

# DEVIATION INEQUALITIES AND MODERATE DEVIATIONS FOR ESTIMATORS OF PARAMETERS IN BIFURCATING AUTOREGRESSIVE MODELS

S.VALÈRE BITSEKI PENDA AND HACÈNE DJELLOUT

**ABSTRACT.** The purpose of this paper is to investigate the deviation inequalities and the moderate deviation principle of the least squares estimators of the unknown parameters of general  $p$ th-order bifurcating autoregressive processes, under suitable assumptions on the driven noise of the process. Our investigation relies on the moderate deviation principle for martingales.

*AMS 2000 subject classifications:* 60F10, 62F12, 60G42, 62M10, 62G05.

## 1. MOTIVATION AND CONTEXT

Bifurcating autoregressive processes (BAR, for short) are an adaptation of autoregressive processes, when the data have a binary tree structure. They were first introduced by Cowan and Staudte [10] for cell lineage data where each individual in one generation gives rise to two offspring in the next generation.

In their paper, the original BAR process was defined as follows. The initial cell is labelled 1, and the two offspring of cell  $k$  are labelled  $2k$  and  $2k+1$ . If  $X_k$  denotes an observation of some characteristic of individual  $k$  then the first order BAR process is given, for all  $k \geq 1$ , by

$$\begin{cases} X_{2k} = a + bX_k + \varepsilon_{2k} \\ X_{2k+1} = a + bX_k + \varepsilon_{2k+1}. \end{cases}$$

The noise sequence  $(\varepsilon_{2k}, \varepsilon_{2k+1})$  represents environmental effects, while  $a, b$  are unknown real parameters, with  $|b| < 1$ , related to inherited effects. The driven noise  $(\varepsilon_{2k}, \varepsilon_{2k+1})$  was originally supposed to be independent and identically distributed with normal distribution. But since two sister cells are in the same environment at their birth,  $\varepsilon_{2k}$  and  $\varepsilon_{2k+1}$  are allowed to be correlated, inducing a correlation between sister cells, distinct from the correlation inherited from their mother.

Several extensions of the model have been proposed and various estimators are studied in the literature for the unknown parameters, see for instance [1],[2], [3], [4], [5], [6]. See [7] for a relevant references.

Recently, there are many studies of the asymmetric BAR process, that is when the quantitative characteristics of the even and odd sisters are allowed to depend from their mother's through different sets of parameters.

Guyon [19] proposes an interpretation of the asymmetric BAR process as a bifurcating Markov chain, which allows him to derive laws of large numbers and central limit theorems

---

*Date:* March 2, 2013.

*Key words and phrases.* Deviation inequalities, Moderate deviation principle, Bifurcating autoregressive process, Martingale, Limit theorems, Least squares estimation.

for the least squares estimators of the unknown parameters of the process. This Markov chain approach was further developed by Delmas and Marsalle [11], where the cells are allowed to die. They defined the genealogy of the cells through a Galton-Watson process, studying the same model on the Galton Watson tree instead of a binary tree.

Another approach based on martingales theory was proposed by Bercu, de Saporta and Gégout-Petit [7], to sharpen the asymptotic analysis of Guyon under weaker assumptions. It must be pointed out that missing data are not dealt with in this work. To take into account possibly missing data in the estimation procedure de Saporta et al. [24] use a two-type Galton-Watson process to model the genealogy.

Our objective in this paper is to go a step further by

- studying the moderate deviation principle (MDP, for short) of the least squares estimators of the unknown parameters of general  $p$ th-order bifurcating autoregressive processes. More precisely we are interested in the asymptotic estimations of

$$\mathbb{P}\left(\frac{\sqrt{n}}{b_n}(\Theta_n - \Theta) \in A\right)$$

where  $\Theta_n$  denotes the estimator of the unknown parameter of interest  $\Theta$ ,  $A$  is a given domain of deviation,  $(b_n > 0)$  is some sequence denoting the scale of deviation. When  $b_n = 1$  this exactly the estimation of the central limit theorem. When  $b_n = \sqrt{n}$ , it becomes the *large deviation*. And when  $1 \ll b_n \ll \sqrt{n}$ , this is the so called *moderate deviations*. Usually, MDP has a simpler rate function inherited from the approximated Gaussian process, and holds for a larger class of dependent random variables than the LDP.

Though we have not found studies exactly on this question in the literatures, except the recent work of Biteski et al. [9] but technically we are much inspired from two lines of studies

- (1) the work of Bercu et al. [7] on the almost sure convergence of the estimators with the quadratic strong law and the central limit theorem;
  - (2) the works of Dembo [12], and Worms [27], [28], [29] on the one hand, and of the paper of Puhalskii [22] and Djellout [15] on the other hand, about the MDP for martingales.
- giving deviation inequalities for the estimator of bifurcating autoregressive processes, which are important for a rigorous non asymptotic statistical study, i.e. for all  $x > 0$

$$\mathbb{P}(\|\Theta_n - \Theta\| \geq x) \leq e^{-C_n(x)},$$

where  $C_n(x)$  will crucially depends on our set of assumptions. The upper bounds in this inequality hold for arbitrary  $n$  and  $x$  (not a limit relation, unlike the MDP results), hence they are much more practical (in statistics). Deviation inequalities for estimators of the parameters associated with linear regression, autoregressive and branching processes are investigated by Bercu and Touati [8]. In the martingale case, deviation inequalities for self normalized martingale have been developed by de la Peña et al. [23]. We also refer to the work of Ledoux [20] for precise credit and references. This type of inequalities are equally well motivated by theoretical question as by numerous applications in different field including the analysis of algorithms, mathematical physics and empirical processes. For some applications in non asymptotic model selection problem we refer to Massart [21].

This paper is organized as follows. First of all, in Section 2, we introduce the  $\text{BAR}(p)$  model as well as the least square estimators for the parameters of observed  $\text{BAR}(p)$  process and some related notation and hypothesis. In Section 3, we state our main results on the deviation inequalities and MDP of our estimators. The section 4 dedicated to the superexponential convergence of the quadratic variation of the martingale, this section contains exponential inequalities which are crucial for the proof of the deviation inequalities. The proofs of the main results are postponed in section 5.

## 2. NOTATIONS AND HYPOTHESIS

In all the sequel, let  $p \in \mathbb{N}^*$ . We consider the asymmetric  $\text{BAR}(p)$  process given, for all  $n \geq 2^{p-1}$ , by

$$\begin{cases} X_{2n} = a_0 + \sum_{k=1}^p a_k X_{[\frac{n}{2^{k-1}}]} + \varepsilon_{2n} \\ X_{2n+1} = b_0 + \sum_{k=1}^p b_k X_{[\frac{n}{2^{k-1}}]} + \varepsilon_{2n+1}, \end{cases} \quad (2.1)$$

where the notation  $[x]$  stands for the largest integer less than or equal to the real  $x$ . The initial states  $\{X_k, 1 \leq k \leq 2^{p-1} - 1\}$  are the ancestors while  $(\varepsilon_{2n}, \varepsilon_{2n+1})$  is the driven noise of the process. The parameters  $(a_0, a_1, \dots, a_p)$  and  $(b_0, b_1, \dots, b_p)$  are unknown real numbers.

The  $\text{BAR}(p)$  process can be rewritten in the abbreviated vector form given, for all  $n \geq 2^{p-1}$ , by

$$\begin{cases} \mathbb{X}_{2n} = A\mathbb{X}_n + \eta_{2n} \\ \mathbb{X}_{2n+1} = B\mathbb{X}_n + \eta_{2n+1} \end{cases} \quad (2.2)$$

where the regression vector  $\mathbb{X}_n = \left(X_n, X_{[\frac{n}{2}]}, \dots, X_{[\frac{n}{2^{p-1}}]}\right)^t$ ,  $\eta_{2n} = (a_0 + \varepsilon_{2n})e_1$ ,  $\eta_{2n+1} = (b_0 + \varepsilon_{2n+1})e_1$ , with  $e_1 = (1, 0, \dots, 0)^t \in \mathbb{R}^p$ . Moreover,  $A$  and  $B$  are the  $p \times p$  companion matrices

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_p \\ 1 & 0 & \cdots & 0 \\ 0 & . & . & . \\ 0 & . & 1 & . \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 & b_2 & \cdots & b_p \\ 1 & 0 & \cdots & 0 \\ 0 & . & . & . \\ 0 & . & 1 & . \end{pmatrix}.$$

In the sequel, we shall assume that the matrices  $A$  and  $B$  satisfy the contraction property

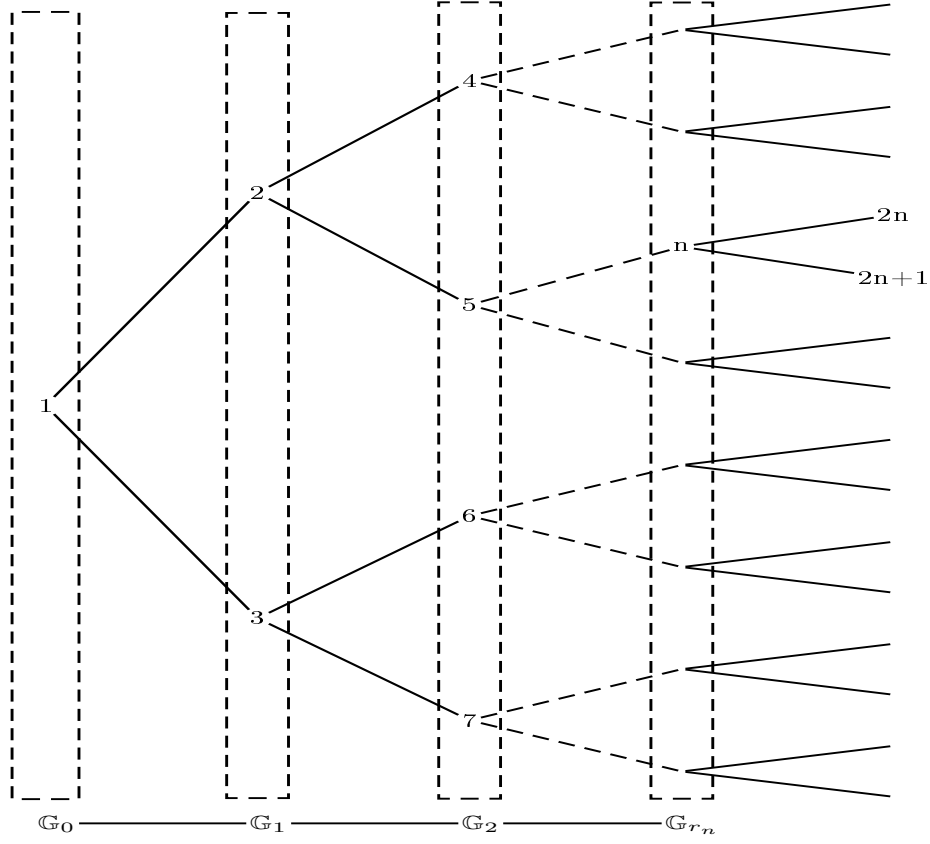
$$\beta = \max(\|A\|, \|B\|) < 1, \quad (2.3)$$

where for any matrix  $M$  the notation  $M^t$ ,  $\|M\|$  and  $\text{Tr}(M)$  stand for the transpose, the euclidean norm and the trace of  $M$ , respectively.

On can see this  $\text{BAR}(p)$  process as a  $p$ th-order autoregressive process on a binary tree, where each vertex represents an individual or cell, vertex 1 being the original ancestor. For all  $n \geq 1$ , denote the  $n$ -th generation by  $\mathbb{G}_n = \{2^n, 2^n + 1, \dots, 2^{n+1} - 1\}$ .

In particular,  $\mathbb{G}_0 = \{1\}$  is the initial generation and  $\mathbb{G}_1 = \{2, 3\}$  is the first generation of offspring from the first ancestor. Let  $\mathbb{G}_{r_n}$  be the generation of individual  $n$ , which means that  $r_n = \lfloor \log_2(n) \rfloor$ . Recall that the two offspring of individual  $n$  are labelled  $2n$  and  $2n+1$ , or conversely, the mother of the individual  $n$  is  $\lfloor n/2 \rfloor$ . More generally, the ancestors of individual  $n$  are  $\lfloor n/2 \rfloor, \lfloor n/2^2 \rfloor, \dots, \lfloor n/2^{r_n} \rfloor$ . Furthermore, denote by

$$\mathbb{T}_n = \bigcup_{k=0}^n \mathbb{G}_k$$

FIGURE 1. The binary tree  $\mathbb{T}$ 

the subtree of all individuals from the original individual up to the  $n$ -th generation. We denote by  $\mathbb{T}_{n,p} = \{k \in \mathbb{T}_n, k \geq 2^p\}$  the subtree of all individuals up to the  $n$ th generation without  $\mathbb{T}_{p-1}$ . One can observe that, for all  $n \geq 1$ ,  $\mathbb{T}_{n,0} = \mathbb{T}_n$  and for all  $p \geq 1$ ,  $\mathbb{T}_{p,p} = \mathbb{G}_p$ .

The  $\text{BAR}(p)$  process can be rewritten, for all  $n \geq 2^{p-1}$ , in the matrix form

$$Z_n = \theta^t Y_n + V_n$$

where

$$Z_n = \begin{pmatrix} X_{2n} \\ X_{2n+1} \end{pmatrix}, \quad Y_n = \begin{pmatrix} 1 \\ \mathbb{X}_n \end{pmatrix}, \quad V_n = \begin{pmatrix} \varepsilon_{2n} \\ \varepsilon_{2n+1} \end{pmatrix},$$

and the  $(p+1) \times 2$  matrix parameter  $\theta$  is given by

$$\theta = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ \cdot & \cdot \\ \cdot & \cdot \\ a_p & b_p \end{pmatrix}.$$

As in Bercu et al.[7], we introduce the least square estimator  $\hat{\theta}_n$  of  $\theta$ , from the observation of all individuals up to the  $n$ -th generation that is the complete sub-tree  $\mathbb{T}_n$ , for all  $n \geq p$

$$\hat{\theta}_n = S_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1,p-1}} Y_k Z_k^t, \quad (2.4)$$

where the  $(p+1) \times (p+1)$  matrix is defined as

$$S_n = \sum_{k \in \mathbb{T}_{n,p-1}} Y_k Y_k^t = \sum_{k \in \mathbb{T}_{n,p-1}} \begin{pmatrix} 1 & \mathbb{X}_k^t \\ \mathbb{X}_k & \mathbb{X}_k \mathbb{X}_k^t \end{pmatrix}. \quad (2.5)$$

We assume, without loss of generality, that for all  $n \geq p-1$ ,  $S_n$  is invertible. In all what follows, we shall make a slight abuse of notation by identifying  $\theta$  as well as  $\hat{\theta}_n$  to

$$\text{vec}(\theta) = \begin{pmatrix} a_0 \\ \cdot \\ \cdot \\ a_p \\ b_0 \\ \cdot \\ \cdot \\ b_p \end{pmatrix} \quad \text{and} \quad \text{vec}(\hat{\theta}_n) = \begin{pmatrix} \hat{a}_{0,n} \\ \cdot \\ \cdot \\ \hat{a}_{p,n} \\ \hat{b}_{0,n} \\ \cdot \\ \cdot \\ \hat{b}_{p,n} \end{pmatrix}.$$

Let  $\Sigma_n = I_2 \otimes S_n$ , where  $\otimes$  stands for the matrix Kronecker product. Therefore, we deduce from (2.4) that

$$\hat{\theta}_n = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1,p-1}} \text{vec}(Y_k Z_k^t) = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1,p-1}} \begin{pmatrix} X_{2k} \\ X_k \mathbb{X}_{2k} \\ X_{2k+1} \\ X_k \mathbb{X}_{2k+1} \end{pmatrix}. \quad (2.6)$$

Consequently, (2.2) yields to

$$\hat{\theta}_n - \theta = \Sigma_{n-1}^{-1} \sum_{k \in \mathbb{T}_{n-1,p-1}} \begin{pmatrix} \varepsilon_{2k} \\ \varepsilon_{2k} \mathbb{X}_k \\ \varepsilon_{2k+1} \\ \varepsilon_{2k+1} \mathbb{X}_k \end{pmatrix}. \quad (2.7)$$

Denote by  $\mathbb{F} = (\mathcal{F}_n)$  the natural filtration associated with the BAR( $p$ ) process, which means that  $\mathcal{F}_n$  is the  $\sigma$ -algebra generated by the individuals up to  $n$ -th generation, in other words  $\mathcal{F}_n = \sigma\{X_k, k \in \mathbb{T}_n\}$ .

For the initial states, if we denote by  $\overline{X}_1 = \max\{\|\mathbb{X}_k\|, k \leq 2^{p-1}\}$ , we introduce the following hypothesis

**(Xa)** For some  $a > 2$ , there exists  $\tau > 0$  such that

$$\mathbb{E} [\exp(\tau \overline{X}_1^a)] < \infty.$$

This assumption implies the weaker Gaussian integrability condition

**(X2)** There is  $\tau > 0$  such that

$$\mathbb{E} [\exp(\tau \overline{X}_1^2)] < \infty.$$

For the noise  $(\varepsilon_{2n}, \varepsilon_{2n+1})$  the assumption may be of two types.

- (1) In the first case we will assume the independence of the noise which allows us to impose less restrictive conditions on the exponential integrability of the noise.

**Case 1:** We shall assume that  $((\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1)$  forms a sequence of independent and identically distributed bi-variate centered random variables with covariance matrix  $\Gamma$  associated with  $(\varepsilon_{2n}, \varepsilon_{2n+1})$ , given by

$$\Gamma = \begin{pmatrix} \sigma^2 & \rho \\ \rho & \sigma^2 \end{pmatrix}, \quad \text{where } \sigma^2 > 0 \text{ and } |\rho| < \sigma^2. \quad (2.8)$$

For all  $n \geq p-1$  and for all  $k \in \mathbb{G}_n$ , we denote

$$\mathbb{E}[\varepsilon_k^2] = \sigma^2, \quad \mathbb{E}[\varepsilon_k^4] = \tau^4, \quad \mathbb{E}[\varepsilon_{2k}\varepsilon_{2k+1}] = \rho, \quad \mathbb{E}[\varepsilon_{2k}^2\varepsilon_{2k+1}^2] = \nu^2 \quad \text{where } \tau^4 > 0, \nu^2 < \tau^4.$$

In addition, we assume that the condition **(X2)** on the initial state is satisfied and

- (G2)** one can find  $\gamma > 0$  and  $c > 0$  such that for all  $n \geq p-1$ , for all  $k \in \mathbb{G}_n$  and for all  $|t| \leq c$

$$\mathbb{E}[\exp t(\varepsilon_k^2 - \sigma^2)] \leq \exp\left(\frac{\gamma t^2}{2}\right).$$

In this case, we impose the following hypothesis on the scale of the deviation

- (V1)**  $(b_n)$  will denote an increasing sequence of positive real numbers such that

$$b_n \longrightarrow +\infty$$

and for  $\beta$  given by (2.3)

- if  $\beta \leq \frac{1}{2}$ , the sequence  $(b_n)$  is such that  $\frac{b_n \log n}{\sqrt{n}} \longrightarrow 0$ ,
- if  $\beta > \frac{1}{2}$ , the sequence  $(b_n)$  is such that  $(b_n \sqrt{\log n})\beta^{\frac{r_{n+1}}{2}} \longrightarrow 0$ .

- (2) In contrast with the first case, in the second case, we will not assume that the sequence  $((\varepsilon_{2n}, \varepsilon_{2n+1}), n \geq 1)$  is i.i.d. The price to pay for giving up this i.i.d. assumption is higher exponential moments. Indeed we need them to make use of the MDP for martingale, especially to prove the Lindeberg condition via Lyapunov one's.

**Case 2:** We shall assume that for all  $n \geq p-1$  and for all  $j \in \mathbb{G}_{n+1}$  that  $\mathbb{E}[\varepsilon_j/\mathcal{F}_n] = 0$  and for all different  $k, l \in \mathbb{G}_{n+1}$  with  $[\frac{k}{2}] \neq [\frac{l}{2}]$ ,  $\varepsilon_k$  and  $\varepsilon_l$  are conditionally independent given  $\mathcal{F}_n$ . And we will use the same notations as in the case 1: for all  $n \geq p-1$  and for all  $k \in \mathbb{G}_{n+1}$

$$\mathbb{E}[\varepsilon_k^2/\mathcal{F}_n] = \sigma^2, \quad \mathbb{E}[\varepsilon_k^4/\mathcal{F}_n] = \tau^4, \quad \mathbb{E}[\varepsilon_{2k}\varepsilon_{2k+1}/\mathcal{F}_n] = \rho, \quad \mathbb{E}[\varepsilon_{2k}^2\varepsilon_{2k+1}^2/\mathcal{F}_n] = \nu^2 \quad a.s.$$

where  $\tau^4 > 0$ ,  $\nu^2 < \tau^4$  and we use also  $\Gamma$  for the conditional covariance matrix associated with  $(\varepsilon_{2n}, \varepsilon_{2n+1})$ . In this case, we assume that the condition **(Xa)** on the initial state is satisfied, and we shall make use of the following hypotheses:

- (Ea)** for some  $a > 2$ , there exist  $t > 0$  and  $E > 0$  such that for all  $n \geq p-1$  and for all  $k \in \mathbb{G}_{n+1}$ ,

$$\mathbb{E}[\exp(t|\varepsilon_k|^{2a})/\mathcal{F}_n] \leq E < \infty \quad a.s.$$

Throughout this case, we introduce the following hypothesis on the scale of the deviation

**(V2)**  $(b_n)$  will denote an increasing sequence of positive real numbers such that

$$b_n \longrightarrow +\infty,$$

and for  $\beta$  given by (2.3)

- if  $\beta^2 < \frac{1}{2}$ , the sequence  $(b_n)$  is such that  $\frac{b_n \log n}{\sqrt{n}} \longrightarrow 0$ ,
- if  $\beta^2 = \frac{1}{2}$ , the sequence  $(b_n)$  is such that  $\frac{b_n (\log n)^{3/2}}{\sqrt{n}} \longrightarrow 0$ ,
- if  $\beta^2 > \frac{1}{2}$ , the sequence  $(b_n)$  is such that  $(b_n \log n) \beta^{r_n+1} \longrightarrow 0$ .

**Remarks 2.1.** *The condition on the scale of the deviation in the case 2, is less restrictive than in the case 1, since we assume more integrability conditions. This condition on the scale of the deviation naturally appear from the calculations (see the proof of Proposition 4.1). Specifically, the log term comes from the crossing of the probability of a sum to the sum of probability.*

**Remarks 2.2.** *From [14] or [20], we deduce with **(Ea)** that*

**(N1)** *there is  $\phi > 0$  such that for all  $n \geq p-1$ , for all  $k \in \mathbb{G}_{n+1}$  and for all  $t \in \mathbb{R}$ ,*

$$\mathbb{E} \left[ \exp(t\varepsilon_k) / \mathcal{F}_n \right] < \exp \left( \frac{\phi t^2}{2} \right), \quad a.s.$$

*We have the same conclusion in the case 1, without the conditioning ; i.e.*

**(G1)** *there is  $\phi > 0$  such that for all  $n \geq p-1$ , for all  $k \in \mathbb{G}_n$  and for all  $t \in \mathbb{R}$ ,*

$$\mathbb{E} \left[ \exp(t\varepsilon_k) \right] < \exp \left( \frac{\phi t^2}{2} \right).$$

**Remarks 2.3.** *Armed by the recent development in the theory of transportation inequalities, exponential integrability and functional inequalities (see Ledoux [20], Gozlan [18] and Gozlan and Leonard [17]), we can prove that a sufficient condition for hypothesis **(G2)** to hold is existence of  $t_0 > 0$  such that for all  $n \geq p-1$  and for all  $k \in \mathbb{G}_n$ ,  $\mathbb{E} [\exp(t_0 \varepsilon_k^2)] < \infty$ .*

We now turn to the estimation of the parameters  $\sigma^2$  and  $\rho$ . On the one hand, we propose to estimate the conditional variance  $\sigma^2$  by

$$\hat{\sigma}_n^2 = \frac{1}{2|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} \|\hat{V}_k\|^2 = \frac{1}{2|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} (\hat{\varepsilon}_{2k}^2 + \hat{\varepsilon}_{2k+1}^2) \quad (2.9)$$

where for all  $n \geq p-1$  and all  $k \in \mathbb{G}_n$ ,  $\hat{V}_k^t = (\hat{\varepsilon}_{2k}, \hat{\varepsilon}_{2k+1})^t$  with

$$\begin{cases} \hat{\varepsilon}_{2k} = X_{2k} - \hat{a}_{0,n} - \sum_{i=1}^p \hat{a}_{i,n} X_{[\frac{k}{2^i-1}]} \\ \hat{\varepsilon}_{2k+1} = X_{2k+1} - \hat{b}_{0,n} - \sum_{i=1}^p \hat{b}_{i,n} X_{[\frac{k}{2^i-1}]} \end{cases}$$

We also introduce the following

$$\sigma_n^2 = \frac{1}{2|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p}} (\varepsilon_{2k}^2 + \varepsilon_{2k+1}^2). \quad (2.10)$$

On the other hand, we estimate the conditional covariance  $\rho$  by

$$\hat{\rho}_n = \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} \hat{\varepsilon}_{2k} \hat{\varepsilon}_{2k+1} \quad (2.11)$$

We also introduce the following

$$\rho_n = \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p}} \varepsilon_{2k} \varepsilon_{2k+1}. \quad (2.12)$$

In order to establish the MDP results of our estimators, we shall make use of a martingale approach. For all  $n \geq p$ , denote

$$M_n = \sum_{k \in \mathbb{T}_{n-1, p-1}} \begin{pmatrix} \varepsilon_{2k} \\ \varepsilon_{2k} \mathbb{X}_k \\ \varepsilon_{2k+1} \\ \varepsilon_{2k+1} \mathbb{X}_k \end{pmatrix} \in \mathbb{R}^{2(p+1)}. \quad (2.13)$$

We can clearly rewrite (2.7) as

$$\hat{\theta}_n - \theta = \Sigma_{n-1}^{-1} M_n. \quad (2.14)$$

We know from Bercu et al. [7] that  $(M_n)$  is a square integrable martingale adapted to the filtration  $\mathbb{F} = (\mathcal{F}_n)$ . Its increasing process is given for all  $n \geq p$  by

$$\langle M \rangle_n = \Gamma \otimes S_{n-1}$$

where  $S_n$  is given in (2.5) and  $\Gamma$  is given in (2.8).

We recall that for a sequence of random variables  $(Z_n)_n$  on  $\mathbb{R}^{d \times p}$ , we say that  $(Z_n)_n$  converges  $(b_n^2)$ -superexponentially fast in probability to some random variable  $Z$  if, for all  $\delta > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P}(\|Z_n - Z\| > \delta) = -\infty.$$

This exponential convergence with speed  $b_n^2$  will be shortened as

$$Z_n \xrightarrow[b_n^2]{\text{superexp}} Z.$$

We follow Dembo and Zeitouni [13] for the language of the large deviations, throughout this paper. Before going further, let us recall the definition of a MDP: let  $(b_n)$  an increasing sequence of positive real numbers such that

$$b_n \longrightarrow \infty \quad \text{and} \quad \frac{b_n}{\sqrt{n}} \longrightarrow 0. \quad (2.15)$$

We say that a sequence of centered random variables  $(M_n)_n$  with topological state space  $(S, \mathcal{S})$  satisfies a MDP with speed  $b_n^2$  and rate function  $I : S \rightarrow \mathbb{R}_+^*$  if for each  $A \in \mathcal{S}$ ,

$$-\inf_{x \in A^\circ} I(x) \leq \liminf_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{\sqrt{n}}{b_n} M_n \in A\right) \leq \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P}\left(\frac{\sqrt{n}}{b_n} M_n \in A\right) \leq -\inf_{x \in \bar{A}} I(x),$$

here  $A^\circ$  and  $\bar{A}$  denote the interior and closure of  $A$  respectively.



Before the presentation of the main results, let us fix some more notation. Let  $\bar{a} = \frac{a_0 + b_0}{2}$ ,  $\bar{a}^2 = \frac{a_0^2 + b_0^2}{2}$ ,  $\bar{A} = \frac{A + B}{2}$  and  $e_1 = (1, 0, \dots, 0)^t \in \mathbb{R}^p$ . We denote

$$\Xi = \bar{a}(I_p - \bar{A})^{-1}e_1, \quad (2.16)$$

and  $\Lambda$  the unique solution of the equation

$$\Lambda = T + \frac{1}{2}(A\Lambda A^t + B\Lambda B^t) \quad (2.17)$$

where

$$T = \left(\sigma^2 + \bar{a}^2\right) e_1 e_1^t + \frac{1}{2} \left(a_0 (A\Xi e_1^t + e_1 \Xi^t A^t) + b_0 (B\Xi e_1^t + e_1 \Xi^t B^t)\right), \quad (2.18)$$

We also introduce the following matrix  $L$  and  $\Sigma$  given by

$$L = \begin{pmatrix} 1 & \Xi \\ \Xi & \Lambda \end{pmatrix} \quad \text{and} \quad \Sigma = I_2 \otimes L. \quad (2.19)$$

**Remarks 2.4.** In the special case  $p = 1$ , we have  $\Xi = \frac{\bar{a}}{1 - \bar{b}}$ , and  $\Lambda = \frac{\bar{a}^2 + \sigma^2 + 2\Xi\bar{a}\bar{b}}{1 - \bar{b}^2}$ , where  $\bar{a}\bar{b} = \frac{a_0 a_1 + b_0 b_1}{2}$ ,  $\bar{b} = \frac{a_1 + b_1}{2}$ ,  $\bar{b}^2 = \frac{a_1^2 + b_1^2}{2}$ .

### 3. MAIN RESULTS

Let us present now the main results of this paper. In the following theorem, we will give the deviation inequalities of the estimator of the parameters, 1 useful for non asymptotic statistical studies.

#### Theorem 3.1.

(i) In the case 1, we have for all  $\delta > 0$  and for all  $b > 0$  such that  $b < \|\Sigma\|/(1 + \delta)$

$$\mathbb{P} \left( \|\hat{\theta}_n - \theta\| > \delta \right) \leq \begin{cases} c_1 \exp \left( -\frac{c_2(\delta b)^2}{c_3 + (\delta b)} \frac{2^n}{(n-1)^2} \right) & \text{if } \beta < \frac{1}{2} \\ c_1(n-1) \exp \left( \frac{-c_2(\delta b)^2}{c_3 + (\delta b)} \frac{2^n}{(n-1)^2} \right) & \text{if } \beta = \frac{1}{2} \\ c_1(n-1) \exp \left( \frac{-c_2(\delta b)^2}{c_3 + (\delta b)} \frac{1}{(n-1)\beta^n} \right) & \text{if } \beta > \frac{1}{2}, \end{cases} \quad (3.1)$$

where the constants  $c_1$ ,  $c_2$  and  $c_3$  depend on  $\sigma^2$ ,  $\beta$ ,  $\gamma$  and  $\phi$  and are such that  $c_1, c_2 > 0$ ,  $c_3 \geq 0$ .

(ii) In the case 2, we have for all  $\delta > 0$  and for all  $b > 0$  such that  $b < \|\Sigma\|/(1 + \delta)$

$$\mathbb{P} \left( \|\hat{\theta}_n - \theta\| > \delta \right) \leq \begin{cases} c_1 \exp \left( -\frac{c_2(\delta b)^2}{c_3 + c_4(\delta b)} \frac{2^n}{(n-1)^2} \right) & \text{if } \beta < \frac{\sqrt{2}}{2} \\ c_1 \exp \left( -\frac{c_2(\delta b)^2}{c_3 + c_4(\delta b)} \frac{2^n}{(n-1)^3} \right) & \text{if } \beta = \frac{\sqrt{2}}{2} \\ c_1 \exp \left( -\frac{c_2(\delta b)^2}{c_3 + c_4(\delta b)} \frac{1}{(n-1)^2 \beta^{2n}} \right) & \text{if } \beta > \frac{\sqrt{2}}{2}, \end{cases} \quad (3.2)$$

where the constants  $c_1, c_2, c_3$ , and  $c_4$  depend on  $\sigma^2, \beta, \gamma$  and  $\phi$  and are such that  $c_1, c_2 > 0, c_3, c_4 \geq 0, (c_3, c_4) \neq (0, 0)$ .

**Remarks 3.2.** One can notice that the estimate (3.2) is stronger than estimate (3.1). This is due to the fact that the integrability condition in case 2 is stronger than integrability condition in case 1.

**Remarks 3.3.** The upper bounds in previous theorem holds for arbitrary  $n \geq p - 1$  (not a limit relation, unlike the below results), hence they are much more practical (in non asymptotic statistics).

In the next result, we will present the MDP of the estimator  $\hat{\theta}_n$ .

**Theorem 3.4.** In the case 1 or in the case 2, the sequence  $\left(\sqrt{|\mathbb{T}_{n-1}|}(\hat{\theta}_n - \theta)/b_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$  satisfies the MDP on  $\mathbb{R}^{2(p+1)}$  with speed  $b_{|\mathbb{T}_{n-1}|}^2$  and rate function

$$I_\theta(x) = \sup_{\lambda \in \mathbb{R}^{2(p+1)}} \{\lambda^t x - \lambda(\Gamma \otimes L^{-1})\lambda^t\} = \frac{1}{2}x^t(\Gamma \otimes L^{-1})^{-1}x, \quad (3.3)$$

where  $L$  and  $\Gamma$  are given in (2.19) and (2.8) respectively.

**Remarks 3.5.** Similar results about deviation inequalities and MDP, are already obtained in [9], in a restrictive case of bounded or gaussian noise and when  $p = 1$ , but results therein hold for general Markov models also.

Let us consider now the estimation of the parameter in the noise process.

**Theorem 3.6.** Let  $(b_n)$  an increasing sequence of positive real numbers such that

$$b_n \longrightarrow \infty \quad \text{and} \quad \frac{b_n}{\sqrt{n}} \longrightarrow 0.$$

In the case 1 or in the case 2,

(1) the sequence  $\left(\sqrt{|\mathbb{T}_{n-1}|}(\sigma_n^2 - \sigma^2)/b_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$  satisfies the MDP on  $\mathbb{R}$  with speed  $b_{|\mathbb{T}_{n-1}|}^2$  and rate function

$$I_{\sigma^2}(x) = \frac{x^2}{\tau^4 - 2\sigma^4 + \nu^2}. \quad (3.4)$$

(2) the sequence  $\left(\sqrt{|\mathbb{T}_{n-1}|}(\rho_n - \rho)/b_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$  satisfies the MDP on  $\mathbb{R}$  with speed  $b_{|\mathbb{T}_{n-1}|}^2$  and rate function

$$I_\rho(x) = \frac{x^2}{2(\nu^2 - \rho^2)}. \quad (3.5)$$

**Remarks 3.7.** Note that in this case the MDP holds for all the scale  $(b_n)$  verifying (2.15) without other restriction.

**Remarks 3.8.** It will be more interesting to prove the MDP for  $\left(\sqrt{|\mathbb{T}_{n-1}|}(\hat{\sigma}_n^2 - \sigma^2)/b_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$ , which will be the case if one proves for example that  $\left(\sqrt{|\mathbb{T}_{n-1}|}(\hat{\sigma}_n^2 - \sigma^2)/b_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$  and

$\left(\sqrt{|\mathbb{T}_{n-1}|}(\sigma_n^2 - \sigma^2)/b_{|\mathbb{T}_{n-1}|}\right)_{n \geq 1}$  are exponentially equivalent in the sense of the MDP. This is described by the following convergence

$$\frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}}(\hat{\sigma}_n^2 - \sigma_n^2) \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0.$$

The proof is very technical and very restrictive for the scale of the deviation. Actually we are only able to prove that

$$\hat{\sigma}_n^2 - \sigma_n^2 \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0,$$

this superexponential convergence will be proved in Theorem 3.9.

In the following theorem we will state the superexponential convergence.

**Theorem 3.9.** *In the case 1 or in the case 2, we have*

$$\hat{\sigma}_n^2 \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} \sigma^2.$$

In the case 1, instead of **(G2)**, if we assume that

**(G2')** one can find  $\gamma' > 0$  such that for all  $n \geq p - 1$ , for all  $k, l \in \mathbb{G}_{n+1}$  with  $[\frac{k}{2}] = [\frac{l}{2}]$  and for all  $t \in ]-c, c[$  for some  $c > 0$ ,

$$\mathbb{E}[\exp t(\varepsilon_k \varepsilon_l - \rho)] \leq \exp\left(\frac{\gamma' t^2}{2}\right),$$

and in the case 2, instead of **(Ea)**, if we assume that

**(E2')** one can find  $\gamma' > 0$  such that for all  $n \geq p - 1$ , for all  $k, l \in \mathbb{G}_{n+1}$  with  $[\frac{k}{2}] = [\frac{l}{2}]$  and for all  $t \in \mathbb{R}$

$$\mathbb{E}[\exp t(\varepsilon_k \varepsilon_l - \rho) / \mathcal{F}_n] \leq \exp\left(\frac{\gamma' t^2}{2}\right), \quad a.s.$$

then in the case 1 or in the case 2, we have

$$\hat{\rho}_n \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} \rho.$$

Before going to the proofs, let us gather here for the convenience of the readers two Theorems useful to establish MDP of the martingales and used intensively in this paper. From this two theorems, we will be able to give a strategy for the proof.

Let  $M = (M_n, \mathcal{H}_n, n \geq 0)$  be a centered square integrable martingale defined on a probability space  $(\Omega, \mathcal{H}, \mathbb{P})$  and  $(\langle M \rangle_n)$  its bracket. Let  $(b_n)$  an increasing sequence of real numbers satisfying (2.15). Let us enunciate the following which corresponds to the unidimensional case of Theorem 1 in [15].

**Proposition 3.10.** *Let  $c(n) := \frac{\sqrt{n}}{b_n}$  is non-decreasing, and define the reciprocal function  $c^{-1}(t)$  by*

$$c^{-1}(t) := \inf\{n \in \mathbb{N} : c(n) \geq t\}.$$

*Under the following conditions:*

**(D1)** *there exists  $Q \in \mathbb{R}_+^*$  such that  $\frac{\langle M \rangle_n}{n} \xrightarrow[b_n^2]{\text{superexp}} Q$ ;*

$$(\mathbf{D2}) \limsup_{n \rightarrow +\infty} \frac{n}{b_n^2} \log \left( n \operatorname{ess\,sup}_{1 \leq k \leq c^{-1}(\sqrt{n+1}b_{n+1})} \mathbb{P}(|M_k - M_{k-1}| > b_n \sqrt{n}/\mathcal{H}_{k-1}) \right) = -\infty;$$

$$(\mathbf{D3}) \text{ for all } a > 0 \frac{1}{n} \sum_{k=1}^n \mathbb{E} \left( |M_k - M_{k-1}|^2 \mathbf{1}_{\{|M_k - M_{k-1}| \geq a \frac{\sqrt{n}}{b_n}\}} / \mathcal{H}_{k-1} \right) \xrightarrow[b_n^2]{\text{superexp}} 0;$$

$(M_n/b_n \sqrt{n})_{n \in \mathbb{N}}$  satisfies the MDP in  $\mathbb{R}$  with the speed  $b_n^2$  and the rate function  $I(x) = \frac{x^2}{2Q}$ .

Let us introduce a simplified version of Puhalskii's result [22] applied to a sequence of martingale differences.

**Theorem 3.11.** *Let  $(m_j^n)_{1 \leq j \leq n}$  be a triangular array of martingale differences with values in  $\mathbb{R}^d$ , with respect to the filtration  $(\mathcal{H}_n)_{n \geq 1}$ . Under the following conditions*

(P1) *there exists a symmetric positive semi-definite matrix  $Q$  such that*

$$\frac{1}{n} \sum_{k=1}^n \mathbb{E} \left[ m_k^n (m_k^n)' | \mathcal{H}_{k-1} \right] \xrightarrow[b_n^2]{\text{superexp}} Q,$$

(P2) *there exists a constant  $c > 0$  such that, for each  $1 \leq k \leq n$ ,  $|m_k^n| \leq c \frac{\sqrt{n}}{b_n}$  a.s.,*

(P3) *for all  $a > 0$ , we have the exponential Lindeberg's condition*

$$\frac{1}{n} \sum_{k=1}^n \mathbb{E} \left[ |m_k^n|^2 \mathbf{1}_{\{|m_k^n| \geq a \frac{\sqrt{n}}{b_n}\}} | \mathcal{H}_{k-1} \right] \xrightarrow[b_n^2]{\text{superexp}} 0.$$

$(\sum_{k=1}^n m_k^n / (b_n \sqrt{n}))_{n \geq 1}$  satisfies an MDP on  $\mathbb{R}^d$  with speed  $b_n^2$  and rate function

$$\Lambda^*(v) = \sup_{\lambda \in \mathbb{R}^d} \left( \lambda' v - \frac{1}{2} \lambda' Q \lambda \right).$$

In particular, if  $Q$  is invertible,  $\Lambda^*(v) = \frac{1}{2} v' Q^{-1} v$ .

As the reader can imagine naturally now, the strategy of the proof of the MDP consist on the following steps :

- the superexponential convergence of the quadratic variation of the martingale  $(M_n)$ . This step is very crucial and the key for the rest of the paper. It will be realized by means of powerful exponential inequalities. This allows us to obtain the deviation inequalities for the estimator of the parameters,
- introduce a truncated martingale which satisfies the MDP, thanks to a classical theorems 3.11,
- the truncated martingale is an exponentially good approximation of  $(M_n)$ , in the sense of the moderate deviation.

#### 4. SUPEREXPONENTIAL CONVERGENCE OF THE QUADRATIC VARIATION OF THE MARTINGALE

At first, it is necessary to establish the superexponential convergence of the quadratic variation of the martingale  $(M_n)$ , properly normalized in order to prove the MDP, of the estimators. Its proof is very technical, but crucial for the rest of the paper. This section contains also some deviation inequalities for some quantities needed in the proof later.

**Proposition 4.1.** *In the case 1 or case 2, we have*

$$\frac{S_n}{|\mathbb{T}_n|} \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} L, \quad (4.1)$$

where  $S_n$  is given in (2.5) and  $L$  is given in (2.19).

For the proof we focus in the case 2. The Proposition 4.1 will follow from Proposition 4.3 and Proposition 4.4 below, where we assume that the sequence  $(b_n)$  satisfies the condition **(V2)**. Proposition 4.10 gives some ideas of the proof in the case 1.

**Remarks 4.2.** *Using [14], we infer from **(Ea)** that*

**(N2)** *one can find  $\gamma > 0$  such that for all  $n \geq p-1$ , for all  $k \in \mathbb{G}_{n+1}$  and for all  $t \in \mathbb{R}$*

$$\mathbb{E} [\exp t (\varepsilon_k^2 - \sigma^2) / \mathcal{F}_n] \leq \exp \left( \frac{\gamma t^2}{2} \right) \quad a.s.$$

**Proposition 4.3.** *Assume that hypothesis **(N2)** and **(Xa)** are satisfied. Then we have*

$$\frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_{n,p}} \mathbb{X}_k \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} \Xi,$$

where  $\Xi$  is given in (2.16).

*Proof.* Let

$$H_n = \sum_{k \in \mathbb{T}_{n,p-1}} \mathbb{X}_k \quad \text{and} \quad P_n = \sum_{k \in \mathbb{T}_{n,p}} \epsilon_k.$$

From Bercu et al. [7], we have

$$\frac{H_n}{2^{n+1}} = \sum_{k=p-1}^n (\bar{A})^{n-k} \frac{H_{p-1}}{2^k} + \sum_{k=p}^n \bar{a}(\bar{A})^{n-k} \left( \frac{2^k - 2^{p-1}}{2^k} \right) e_1 + \sum_{k=p}^n \frac{P_k}{2^{k+1}} (\bar{A})^{n-k} e_1. \quad (4.2)$$

Since the second term in the right hand side of this equality is deterministic, this proposition will be proved if we show that

$$\sum_{k=p-1}^n \frac{(\bar{A})^{n-k}}{2^k} H_{p-1} \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} 0, \quad \sum_{k=p}^n \frac{P_k}{2^{k+1}} (\bar{A})^{n-k} e_1 \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} 0, \quad (4.3)$$

which follows by performing as in the proof of Proposition 4.4 (see the proof of Proposition 4.4 for more details).  $\square$

**Proposition 4.4.** *Assume that hypothesis **(N2)** and **(Xa)** are satisfied. Then we have*

$$\frac{1}{|\mathbb{T}_n|} \sum_{k \in \mathbb{T}_{n,p}} \mathbb{X}_k \mathbb{X}_k^t \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} \Lambda,$$

where  $\Lambda$  is given in (2.17).

*Proof.* Let

$$K_n = \sum_{k \in \mathbb{T}_{n,p-1}} \mathbb{X}_k \mathbb{X}_k^t \quad \text{and} \quad L_n = \sum_{k \in \mathbb{T}_{n,p}} \varepsilon_k^2. \quad (4.4)$$

Then from (2.2), and after straightforward calculations (see [7] for more details), we get that

$$\frac{K_n}{2^{n+1}} = \frac{1}{2^{n-p+1}} \sum_{C \in \{A;B\}^{n-p+1}} C \frac{K_{p-1}}{2^p} C^t + \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C T_{n-k} C^t,$$

where the notation  $\{A;B\}^k$  means the set of all products of  $A$  and  $B$  with exactly  $k$  terms. The cardinality of  $\{A;B\}^k$  is obviously  $2^k$ , and

$$T_k = \frac{L_k}{2^{k+1}} e_1 e_1^t + \overline{a^2} \left( \frac{2^k - 2^{p-1}}{2^k} \right) e_1 e_1^t + I_k^{(1)} + I_k^{(2)} + \frac{1}{2^{k+1}} U_k$$

with  $\overline{a^2} = (a_0^2 + b_0^2)/2$  and

$$I_k^{(1)} = \frac{1}{2} \left( a_0 \left( A \frac{H_{k-1}}{2^k} e_1^t + e_1 \frac{H_{k-1}}{2^k} A^t \right) + b_0 \left( B \frac{H_{k-1}}{2^k} e_1^t + e_1 \frac{H_{k-1}}{2^k} B^t \right) \right), \quad (4.5)$$

$$I_k^{(2)} = \left( \frac{1}{2^k} \sum_{l \in \mathbb{T}_{k-1,p-1}} (a_0 \varepsilon_{2l} + b_0 \varepsilon_{2l+1}) \right) e_1 e_1^t, \quad (4.6)$$

$$U_k = \sum_{l \in \mathbb{T}_{k-1,p-1}} \varepsilon_{2l} \left( A \mathbb{X}_l e_1^t + e_1 \mathbb{X}_l^t A^t \right) + \varepsilon_{2l+1} \left( B \mathbb{X}_l e_1^t + e_1 \mathbb{X}_l^t B^t \right). \quad (4.7)$$

Then proposition will follow if we prove Lemmas 4.5, 4.6, 4.7, 4.8 and 4.9.

**Lemma 4.5.** *Assume that hypothesis (Xa) is satisfied. Then we have*

$$\frac{1}{2^{n-p+1}} \sum_{C \in \{A;B\}^{n-p+1}} C \frac{K_{p-1}}{2^p} C^t \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} 0, \quad (4.8)$$

where  $K_p$  is given in (4.4).

*Proof.* We get easily

$$\left\| \frac{1}{2^{n-p+1}} \sum_{C \in \{A;B\}^{n-p+1}} C \frac{K_{p-1}}{2^p} C^t \right\| \leq c \beta^{2n} \overline{X}_1^2,$$

where  $\beta$  is given in (2.3),  $\overline{X}_1$  is introduced in (Xa) and  $c$  is a positive constant which depends on  $p$ . Next, Chernoff inequality and hypothesis (X2) lead us easily to (4.8).  $\square$

**Lemma 4.6.** *Assume that hypothesis (N2) and (Xa) are satisfied. Then we have*

$$\sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C \frac{L_{n-k}}{2^{n-k}} e_1 e_1^t C^t \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} \sigma^2 e_1 e_1^t, \quad (4.9)$$

where  $L_k$  is given in the second part of (4.4).

*Proof.* First, since we have for all  $k \geq p$  the following decomposition on odd and even part

$$\sum_{i \in \mathbb{T}_{k,p}} (\varepsilon_i^2 - \sigma^2) = \sum_{i \in \mathbb{T}_{k-1,p-1}} (\varepsilon_{2i}^2 - \sigma^2) + (\varepsilon_{2i+1}^2 - \sigma^2),$$

we obtain for all  $\delta > 0$  that

$$\mathbb{P} \left( \frac{1}{|\mathbb{T}_k| + 1} \sum_{i \in \mathbb{T}_{k,p}} (\varepsilon_i^2 - \sigma^2) > \delta \right) \leq \sum_{\eta=0}^1 \mathbb{P} \left( \frac{1}{|\mathbb{T}_k| + 1} \sum_{i \in \mathbb{T}_{k-1,p-1}} (\varepsilon_{2i+\eta}^2 - \sigma^2) > \frac{\delta}{2} \right).$$

We will treat only the case  $\eta = 0$ . Chernoff inequality gives us for all  $\lambda > 0$

$$\mathbb{P} \left( \frac{1}{|\mathbb{T}_k| + 1} \sum_{i \in \mathbb{T}_{k-1,p-1}} (\varepsilon_{2i}^2 - \sigma^2) > \frac{\delta}{2} \right) \leq \exp \left( -\lambda \frac{\delta}{2} 2^{k+1} \right) \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{T}_{k-1,p-1}} (\varepsilon_{2i}^2 - \sigma^2) \right) \right].$$

We obtain from hypothesis **(N2)**, after conditioning by  $\mathcal{F}_{k-1}$

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{T}_{k-1,p-1}} (\varepsilon_{2i}^2 - \sigma^2) \right) \right] \leq \exp (\lambda^2 \gamma |\mathbb{G}_{k-1}|) \mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{T}_{k-2,p-1}} (\varepsilon_{2i}^2 - \sigma^2) \right) \right].$$

Iterating this, we deduce that

$$\mathbb{E} \left[ \exp \left( \lambda \sum_{i \in \mathbb{T}_{k-1,p-1}} (\varepsilon_{2i}^2 - \sigma^2) \right) \right] \leq \exp \left( \gamma \lambda^2 \sum_{l=p-1}^{k-1} |\mathbb{G}_l| \right) \leq \exp (\gamma \lambda^2 2^{k+1}).$$

Next, optimizing on  $\lambda$ , we get

$$\mathbb{P} \left( \frac{1}{|\mathbb{T}_k| + 1} \sum_{i \in \mathbb{T}_{k-1,p-1}} (\varepsilon_{2i}^2 - \sigma^2) > \frac{\delta}{2} \right) \leq \exp (-c\delta^2 |\mathbb{T}_k|)$$

for some positive constant  $c$  which depends on  $\gamma$ . Applying the foregoing to the random variables  $-(\varepsilon_i^2 - \sigma^2)$ , we obtain

$$\mathbb{P} \left( \frac{1}{|\mathbb{T}_k| + 1} \left| \sum_{i \in \mathbb{T}_{k,p}} (\varepsilon_i^2 - \sigma^2) \right| > \delta \right) \leq 4 \exp (-c\delta^2 |\mathbb{T}_k|). \quad (4.10)$$

Next, from the following inequalities

$$\begin{aligned} \left\| \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C \frac{L_{n-k} - \sigma^2}{2^{n-k}} e_1 e_1^t C^t \right\| &\leq \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} \frac{|L_{n-k} - \sigma^2|}{2^{n-k}} \|C e_1 e_1^t C^t\| \\ &\leq \sum_{k=p}^n \beta^{2(n-k)} \frac{|L_k - \sigma^2|}{|\mathbb{T}_k| + 1} \end{aligned}$$

and from (4.10) applied with  $\delta/((n-p+1)\beta^{2(n-k)})$  instead of  $\delta$ , we get

$$\begin{aligned} \mathbb{P} \left( \left\| \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C \frac{L_{n-k} - \sigma^2}{2^{n-k}} e_1 e_1^t C^t \right\| > \delta \right) &\leq \mathbb{P} \left( \sum_{k=p}^n \beta^{2(n-k)} \frac{|L_k - \sigma^2|}{|\mathbb{T}_k| + 1} > \delta \right) \\ &\leq \sum_{k=p}^n \mathbb{P} \left( \frac{|L_k - \sigma^2|}{|\mathbb{T}_k| + 1} > \frac{\delta}{(n-p+1)\beta^{2(n-k)}} \right) \\ &\leq c_1 \sum_{k=p}^n \exp \left( -c_2 \delta^2 \frac{(2\beta^4)^{k+1}}{n^2 \beta^{4n}} \right). \end{aligned}$$

Now, following the same lines as in the proof of (4.17) we obtain

$$\mathbb{P} \left( \left\| \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C \frac{L_{n-k} - \sigma^2}{2^{n-k}} e_1 e_1^t C^t \right\| > \delta \right) \leq \begin{cases} c_1 \exp \left( -c_2 \delta^2 \frac{2^{n+1}}{n^2} \right) & \text{if } \beta^4 < \frac{1}{2}, \\ c_1 n \exp \left( -c_2 \delta^2 \frac{2^{n+1}}{n^2} \right) & \text{if } \beta^4 = \frac{1}{2}, \\ c_1 \exp \left( -c_2 \delta^2 \frac{1}{n^2 \beta^{4n}} \right) & \text{if } \beta^4 > \frac{1}{2}, \end{cases} \quad (4.11)$$

for some positive constants  $c_1$  and  $c_2$ . From (4.11), we infer that (4.9) holds.  $\square$

**Lemma 4.7.** *Assume that hypothesis (N2) is satisfied. Then we have*

$$\sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C I_{n-k}^{(2)} C^t \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} 0, \quad (4.12)$$

where  $I_k^{(2)}$  is given in (4.6).

*Proof.* This proof follows the same lines as that of (4.9) and uses hypothesis (N1) instead of (N2).  $\square$

**Lemma 4.8.** *Assume that hypothesis (N2) and (Xa) are satisfied. Then we have*

$$\sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C I_{n-k}^{(1)} C^t \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} \Lambda', \quad \text{where } \Lambda' = T - (\sigma^2 + \overline{a^2}) e_1 e_1^t, \quad (4.13)$$

where  $T$  is given (2.18) and  $I_k^{(1)}$  is given in (4.5).

*Proof.* Since in the definition of  $I_n^{(1)}$  given by (4.5), there are four terms, we will focus only on the first term

$$\frac{a_0}{2} A \frac{H_{k-1}}{2^k} e_1^t,$$

the other terms will be treated in the same way. Using (4.2), we obtain the following decomposition:

$$\frac{a_0}{2} \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C A \frac{H_{n-k-1}}{2^{n-k}} e_1^t C^t = T_n^{(1)} + T_n^{(2)} + T_n^{(3)}$$



where

$$T_n^{(1)} = \frac{a_0}{2} \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} CA \left\{ \bar{A}^{n-k-p} \frac{H_{p-1}}{2^p} + \sum_{l=p}^{n-k-1} \bar{A}^{n-k-l-1} \frac{H_{p-1}}{2^{l+1}} \right\} e_1^t C^t,$$

$$T_n^{(2)} = \frac{a_0}{2} \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} CA \left\{ \sum_{l=p}^{n-k-1} \bar{A}^{n-k-l-1} \bar{a} \left( \frac{2^l - 2^{p-1}}{2^l} \right) e_1 e_1^t \right\} C^t,$$

and

$$T_n^{(3)} = \frac{a_0}{2} \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} CA \sum_{l=p}^{n-k-1} \bar{A}^{n-k-l-1} \frac{P_l}{2^{l+1}} e_1 e_1^t C^t.$$

On the one hand we have

$$\|T_n^{(3)}\| \leq c \sum_{k=p}^n \beta^{n-k} \frac{|P_k|}{2^{k+1}}$$

where  $c$  is a positive constant such that  $c > |a_0| \frac{1-\beta^{n-l}}{1-\beta}$  for all  $n \geq l$ , so that

$$\mathbb{P}(\|T_n^{(3)}\| > \delta) \leq \sum_{k=p}^n \mathbb{P}\left(\frac{|P_k|}{|\mathbb{T}_k| + 1} > \frac{2\delta}{cn\beta^{n-k}}\right).$$

We deduce again from hypothesis **(N1)** and in the same way we have obtained (4.10) that

$$\mathbb{P}\left(\frac{|P_k|}{|\mathbb{T}_k| + 1} > \frac{2\delta}{cn\beta^{n-k}}\right) \leq \exp\left(-c_1 \delta^2 \frac{(2\beta^2)^{k+1}}{n^2 \beta^{2n}}\right) \quad \forall k \geq p,$$

for some positive constant  $c_1$ . It then follows as in the proof of (4.17) that

$$\mathbb{P}(\|T_n^{(3)}\| > \delta) \leq \begin{cases} \exp\left(-c_1 \delta^2 \frac{2^{n+1}}{n^2}\right) & \text{if } \beta^2 < \frac{1}{2}, \\ n \exp\left(-c_1 \delta^2 \frac{2^{n+1}}{n^2}\right) & \text{if } \beta^2 = \frac{1}{2}, \\ \exp\left(-c_1 \delta^2 \frac{1}{n^2 \beta^{2n}}\right) & \text{if } \beta^2 > \frac{1}{2}, \end{cases}$$

so that

$$T_n^{(3)} \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} 0. \quad (4.14)$$

On the other hand, we have after studious calculations

$$\|T_n^{(1)}\| \leq \begin{cases} c \frac{\bar{X}_1}{2^{n+1}} & \text{if } \beta < \frac{1}{2}, \\ c \frac{\bar{X}_1}{\sqrt{|\mathbb{T}_n|+1}} & \text{if } \beta = \frac{1}{2}, \\ c \beta^n \bar{X}_1 & \text{if } \beta > \frac{1}{2}, \end{cases}$$

where  $c$  is a positive constant which depends on  $p$  and  $|a_0|$ . Next, from hypothesis **(X2)** and Chernoff inequality we conclude that

$$T_n^{(1)} \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} 0. \quad (4.15)$$

Furthermore, since  $(T_n^{(2)})$  is a deterministic sequence, we have

$$T_n^{(2)} \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} \frac{1}{2} a_0 A \Xi e_1^t. \quad (4.16)$$

It then follows that

$$\frac{a_0}{2} \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C A \frac{H_{n-k-1}}{2^{n-k}} e_1^t C^t \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} \frac{1}{2} a_0 A \Xi e_1^t.$$

Doing the same for the three other terms of  $I_k^{(1)}$ , we end the proof of Lemma (4.8).  $\square$

**Lemma 4.9.** *Assume that hypothesis (N2) and (Xa) are satisfied. Then we have*

$$\sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C \frac{U_{n-k}}{2^{n-k+1}} C^t \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} 0, \quad (4.17)$$

where  $U_k$  is given by (4.7).

*Proof.* Let  $V_n = \sum_{k=2^{p-1}}^n \varepsilon_{2k} X_k$ . Then  $(V_n)$  is a  $\mathcal{G}_n$ -martingale and its increasing process verifies that

$$\langle V \rangle_n = \sigma^2 \sum_{k=2^{p-1}}^n X_k^2 \leq \sigma^2 \sum_{k=2^{p-1}}^n \|\mathbb{X}_k\|^2 \leq \sigma^2 \sum_{k \in \mathbb{T}_{r_n, p-1}} \|\mathbb{X}_k\|^2$$

From [7], with  $\alpha = \max(|a_0|, |b_0|)$ , we have

$$\sum_{k \in \mathbb{T}_{r_n, p-1}} \|\mathbb{X}_k\|^2 \leq \frac{4}{1-\beta} P_{r_n} + \frac{4\alpha^2}{1-\beta} Q_{r_n} + 2\bar{X}_1^2 R_{r_n}, \quad (4.18)$$

where

$$P_{r_n} = \sum_{k \in \mathbb{T}_{r_n, p}} \sum_{i=0}^{r_k-p} \beta^i \varepsilon_{[\frac{k}{2^i}]}^2, \quad Q_{r_n} = \sum_{k \in \mathbb{T}_{r_n, p}} \sum_{i=0}^{r_k-p} \beta^i, \quad R_{r_n} = \sum_{k \in \mathbb{T}_{r_n, p-1}} \beta^{2(r_k-p+1)}.$$

For  $\lambda > 0$ , we infer from hypothesis (N1) that  $(Y_k)_{2^{p-1} \leq k \leq n}$  given by

$$Y_n = \exp \left( \lambda V_n - \frac{\lambda^2 \phi}{2} \sum_{k=2^{p-1}}^n X_k^2 \right),$$

is a  $\mathcal{G}_k$ -supermartingale and moreover  $\mathbb{E}[Y_{2^{p-1}}] \leq 1$ .

For  $B > 0$  and  $\delta > 0$ , we have

$$\begin{aligned} \mathbb{P} \left( \frac{V_n}{2n} > \delta \right) &\leq \mathbb{P} \left( \frac{\phi}{2n} \sum_{i=2^{p-1}}^n X_k^2 > B \right) + \mathbb{P} \left( Y_n > \exp \left( \lambda \delta - \frac{\lambda^2 B}{2} \right) 2n \right) \\ &\leq \mathbb{P} \left( \frac{\phi}{2n} \sum_{k=2^{p-1}}^n X_k^2 > B \right) + \exp \left( \left( -\lambda \delta + \frac{\lambda^2 B}{2} \right) 2n \right). \end{aligned}$$

Optimizing on  $\lambda$ , we get

$$\mathbb{P} \left( \frac{V_n}{2n} > \delta \right) \leq \mathbb{P} \left( \frac{\phi}{2n} \sum_{k \in \mathbb{T}_{r_n, p-1}} \|\mathbb{X}_k\|^2 > B \right) + \exp \left( -\frac{\delta^2}{B} 2n \right).$$

Since the same thing works for  $-V_n$  instead of  $V_n$ , using  $|\mathbb{T}_{n-1}|$  instead of  $n$  in the previous inequality, we have particularly

$$\mathbb{P} \left( \frac{|V_{|\mathbb{T}_{n-1}|}|}{|\mathbb{T}_n| + 1} > \delta \right) \leq \mathbb{P} \left( \frac{\phi}{|\mathbb{T}_n| + 1} \sum_{k \in \mathbb{T}_{n-1, p-1}} \|\mathbb{X}_k\|^2 > B \right) + \exp \left( -\frac{\delta^2}{B} 2^{n+1} \right). \quad (4.19)$$

Now, to control the first term in the right hand of the last inequality, we will use the decomposition given by (4.18). From the convergence of  $\frac{4\phi}{(1-\beta)(|\mathbb{T}_n|+1)}P_n$  and  $\frac{4\phi\alpha^2}{(1-\beta)(|\mathbb{T}_n|+1)}Q_n$  (see [7] for more details) let  $l_1$  and  $l_2$  such that  $\forall n \geq p-1$

$$\frac{4\phi P_{n-1}}{(1-\beta)(|\mathbb{T}_n|+1)} \rightarrow l_1 \quad \text{and} \quad \frac{4\phi\alpha^2 Q_{n-1}}{(1-\beta)(|\mathbb{T}_n|+1)} < l_2.$$

For  $\delta > 0$ , we choose  $B = \delta + l_1 + l_2$ , using (4.18), we then have

$$\begin{aligned} & \mathbb{P} \left( \frac{\phi}{|\mathbb{T}_n| + 1} \sum_{k \in \mathbb{T}_{n-1, p-1}} \|\mathbb{X}_k\|^2 > B \right) \\ & \leq \mathbb{P} \left( \frac{P_{n-1}}{|\mathbb{T}_n| + 1} - l'_1 > \delta_1 \right) + \mathbb{P} \left( \frac{Q_{n-1}}{|\mathbb{T}_n| + 1} - l'_2 > \delta_2 \right) + \mathbb{P} \left( \frac{R_{n-1} \bar{X}_1^2}{|\mathbb{T}_n| + 1} > \delta_3 \right) \end{aligned} \quad (4.20)$$

where

$$\delta_1 = \frac{(1-\beta)\delta}{12\phi}, \quad l'_1 = \frac{(1-\beta)l_1}{4\phi}, \quad \delta_2 = \frac{(1-\beta)\delta}{12\alpha^2\phi}, \quad l'_2 = \frac{(1-\beta)l_2}{4\alpha^2\phi}, \quad \text{and} \quad \delta_3 = \frac{\delta}{6\phi}.$$

First, by the choice of  $l_2$ , we have

$$\mathbb{P} \left( \frac{Q_{n-1}}{|\mathbb{T}_n| + 1} - l'_2 > \delta_2 \right) = 0. \quad (4.21)$$

Next, from Chernoff inequality and hypothesis **(X2)** we get easily

$$\mathbb{P} \left( \frac{R_{n-1} \bar{X}_1^2}{|\mathbb{T}_n| + 1} > \delta_3 \right) \leq \begin{cases} c_1 \exp \left( -c_2 \delta 2^{n+1} \right) & \text{if } \beta < \frac{\sqrt{2}}{2} \\ c_1 \exp \left( -c_2 \delta \frac{2^{n+1}}{n+1} \right) & \text{if } \beta = \frac{\sqrt{2}}{2} \\ c_1 \exp \left( -c_2 \delta \left( \frac{1}{\beta^2} \right)^{n+1} \right) & \text{if } \beta > \frac{\sqrt{2}}{2}, \end{cases} \quad (4.22)$$

for some positive constants  $c_1$  and  $c_2$ . Let us now control the first term of the right hand side of (4.20).

**First case.** If  $\beta = \frac{1}{2}$ , from [7]

$$P_{n-1} = \sum_{k=p}^{n-1} (n-k) \sum_{i \in \mathbb{G}_k} \varepsilon_i^2 \quad \text{and} \quad l'_1 = \sigma^2.$$

We thus have

$$\frac{P_{n-1}}{|\mathbb{T}_n| + 1} - \sigma^2 = \frac{1}{|\mathbb{T}_n| + 1} \sum_{k=p}^{n-1} (n-k) \sum_{i \in \mathbb{G}_k} (\varepsilon_i^2 - \sigma^2) + \sigma^2 \left( \sum_{k=p}^{n-1} \frac{n-k}{2^{n+1-k}} - 1 \right).$$

In addition, we also have

$$\sigma^2 \left( \sum_{k=p}^{n-1} \frac{n-k}{2^{n+1-k}} - 1 \right) \leq 0.$$

We thus deduce that

$$\mathbb{P} \left( \frac{P_{n-1}}{|\mathbb{T}_n| + 1} - l'_1 > \delta_1 \right) \leq \mathbb{P} \left( \frac{1}{|\mathbb{T}_n| + 1} \sum_{k=p}^{n-1} (n-k) \sum_{i \in \mathbb{G}_k} (\varepsilon_i^2 - \sigma^2) > \delta_1 \right).$$

On the one hand we have

$$\begin{aligned} & \mathbb{P} \left( \frac{1}{|\mathbb{T}_n| + 1} \sum_{k=p}^{n-1} (n-k) \sum_{i \in \mathbb{G}_k} (\varepsilon_i^2 - \sigma^2) > \delta_1 \right) \\ & \leq \sum_{\eta=0}^1 \mathbb{P} \left( \frac{1}{|\mathbb{T}_n| + 1} \sum_{k=p-1}^{n-2} (n-k-1) \sum_{i \in \mathbb{G}_k} (\varepsilon_{2i+\eta}^2 - \sigma^2) > \delta_1/2 \right). \end{aligned} \quad (4.23)$$

On the other hand, for all  $\lambda > 0$ , an application of Chernoff inequality yields

$$\begin{aligned} & \mathbb{P} \left( \frac{1}{|\mathbb{T}_n| + 1} \sum_{k=p-1}^{n-2} (n-k-1) \sum_{i \in \mathbb{G}_k} (\varepsilon_{2i}^2 - \sigma^2) > \delta_1/2 \right) \\ & \leq \exp \left( \frac{-\delta_1 \lambda 2^{n+1}}{2} \right) \times \mathbb{E} \left[ \exp \left( \lambda \sum_{k=p-1}^{n-2} (n-k-1) \sum_{i \in \mathbb{G}_k} (\varepsilon_{2i}^2 - \sigma^2) \right) \right]. \end{aligned}$$

From hypothesis **(N2)** we get

$$\begin{aligned} & \mathbb{E} \left[ \exp \left( \lambda \sum_{k=p-1}^{n-2} (n-k-1) \sum_{i \in \mathbb{G}_k} (\varepsilon_{2i}^2 - \sigma^2) \right) \right] \\ & = \mathbb{E} \left[ \mathbb{E} \left[ \exp \left( \lambda \sum_{k=p-1}^{n-2} (n-k-1) \sum_{i \in \mathbb{G}_k} (\varepsilon_{2i}^2 - \sigma^2) \right) / \mathcal{F}_n \right] \right] \\ & = \mathbb{E} \left[ \exp \left( \lambda \sum_{k=p-1}^{n-3} (n-k-1) \sum_{i \in \mathbb{G}_k} (\varepsilon_{2i}^2 - \sigma^2) \right) \prod_{i \in \mathbb{G}_{n-2}} \mathbb{E} \left[ \exp (\lambda (\varepsilon_{2i}^2 - \sigma^2)) / \mathcal{F}_n \right] \right] \\ & \leq \exp (\lambda^2 \gamma |\mathbb{G}_{n-2}|) \mathbb{E} \left[ \exp \left( \lambda \sum_{k=p-1}^{n-3} (n-k-1) \sum_{i \in \mathbb{G}_k} (\varepsilon_{2i}^2 - \sigma^2) \right) \right]. \end{aligned}$$

Iterating this procedure, we obtain

$$\begin{aligned} \mathbb{E} \left[ \exp \left( \lambda \sum_{k=p-1}^{n-2} (n-k-1) \sum_{i \in \mathbb{G}_k} (\varepsilon_{2i}^2 - \sigma^2) \right) \right] & \leq \exp \left( \gamma \lambda^2 \sum_{k=2}^{n-p+1} k^2 |\mathbb{G}_{n-k}| \right) \\ & \leq \exp (c \gamma \lambda^2 2^{n+1}), \end{aligned}$$

where  $c = \sum_{k=1}^{\infty} \frac{k^2}{2^{k+2}}$ . Optimizing on  $\lambda$ , we are led, for some positive constant  $c_1$  to

$$\mathbb{P} \left( \frac{1}{|\mathbb{T}_n| + 1} \sum_{k=p-1}^{n-2} (n-k-1) \sum_{i \in \mathbb{G}_k} (\varepsilon_{2i}^2 - \sigma^2) > \delta_1/2 \right) \leq \exp(-c_1 \delta^2 |\mathbb{T}_n|).$$

Following the same lines, we obtain the same inequality for the second term in (4.23). It then follows that

$$\mathbb{P} \left( \frac{P_{n-1}}{|\mathbb{T}_n| + 1} - l'_1 > \delta_1 \right) \leq c_1 \exp(-c_2 \delta^2 |\mathbb{T}_n|), \quad (4.24)$$

for some positive constants  $c_1$  and  $c_2$ .

**Second case.** If  $\beta \neq \frac{1}{2}$ , then from [7], we have  $l'_1 = \frac{\sigma^2}{2(1-\beta)}$ . Since

$$\sigma^2 \left( \sum_{k=p}^{n-1} \frac{1 - (2\beta)^{n-k}}{(1-2\beta)2^{n-k+1}} \right) \leq \frac{\sigma^2}{2(1-\beta)},$$

we deduce that

$$\mathbb{P} \left( \frac{P_{n-1}}{|\mathbb{T}_n| + 1} - l'_1 > \delta_1 \right) \leq \mathbb{P} \left( \frac{1}{|\mathbb{T}_n| + 1} \sum_{k=p}^{n-1} \frac{1 - (2\beta)^{n-k}}{1 - 2\beta} \sum_{i \in \mathbb{G}_k} (\varepsilon_i^2 - \sigma^2) > \delta_1 \right).$$

- If  $\beta < \frac{1}{2}$ , then for some positive constant  $c$  we have

$$\mathbb{P} \left( \frac{P_{n-1}}{|\mathbb{T}_n| + 1} - l'_1 > \delta_1 \right) \leq \mathbb{P} \left( \frac{1}{|\mathbb{T}_n| + 1} \sum_{k=p}^{n-1} \sum_{i \in \mathbb{G}_k} (\varepsilon_i^2 - \sigma^2) > c\delta_1 \right).$$

Performing now as in the proof of (4.9), we get

$$\mathbb{P} \left( \frac{P_{n-1}}{|\mathbb{T}_n| + 1} - l'_1 > \delta_1 \right) \leq c_1 \exp(-c_2 \delta^2 |\mathbb{T}_n|), \quad (4.25)$$

for some positive constants  $c_1$  and  $c_2$ .

- If  $\beta > \frac{1}{2}$ , then for some positive constant  $c$ , we have

$$\mathbb{P} \left( \frac{P_{n-1}}{|\mathbb{T}_n| + 1} - l'_1 > \delta_1 \right) \leq \mathbb{P} \left( \frac{1}{|\mathbb{T}_n| + 1} \sum_{k=p}^{n-1} (2\beta)^{n-k} \sum_{i \in \mathbb{G}_k} (\varepsilon_i^2 - \sigma^2) > c\delta_1 \right).$$

Now, from Chernoff inequality, hypothesis **(N2)** and after several successive conditioning, we get for all  $\lambda > 0$

$$\begin{aligned} & \mathbb{P} \left( \frac{1}{|\mathbb{T}_n| + 1} \sum_{k=p}^{n-1} (2\beta)^{n-k} \sum_{i \in \mathbb{G}_k} (\varepsilon_i^2 - \sigma^2) > c\delta_1 \right) \\ & \leq \exp(-c\delta_1 \lambda 2^{n+1}) \exp \left( \gamma \lambda^2 2^{n+1} \sum_{k=2}^{n-p+1} (2\beta^2)^k \right). \end{aligned}$$

Next, optimizing over  $\lambda$ , we are led, for some positive constant  $c$  to

$$\mathbb{P} \left( \frac{P_{n-1}}{|\mathbb{T}_n| + 1} - l'_1 > \delta_1 \right) \leq \begin{cases} \exp \left( -c\delta^2 |\mathbb{T}_n| \right) & \text{if } \frac{1}{2} < \beta < \frac{\sqrt{2}}{2}, \\ \exp \left( -c\delta^2 \frac{|\mathbb{T}_n|}{n} \right) & \text{if } \beta = \frac{\sqrt{2}}{2}, \\ \exp \left( -c\delta^2 \left( \frac{1}{\beta^2} \right)^{n+1} \right) & \text{if } \beta > \frac{\sqrt{2}}{2}. \end{cases} \quad (4.26)$$

Now combining (4.19), (4.20), (4.21), (4.22), (4.24), (4.25) and (4.26), we have thus showed that

$$\begin{aligned} & \mathbb{P} \left( \frac{1}{|\mathbb{T}_n| + 1} |V_{|\mathbb{T}_{n-1}|}| > \delta \right) \\ & \leq \begin{cases} c_1 \exp(-c_2 \delta^2 2^{n+1}) + c_1 \exp(-c_2 \delta 2^{n+1}) + \exp \left( \frac{-\delta^2}{\delta + l_1 + l_2} 2^{n+1} \right) & \text{if } \beta < \frac{\sqrt{2}}{2}, \\ c_1 \exp \left( -c_2 \delta^2 \frac{2^{n+1}}{n+1} \right) + c_1 \exp \left( -c_2 \delta \frac{2^{n+1}}{n+1} \right) + \exp \left( \frac{-\delta^2}{\delta + l_1 + l_2} 2^{n+1} \right) & \text{if } \beta = \frac{\sqrt{2}}{2}, \\ c_1 \exp \left( -c_2 \delta^2 \left( \frac{1}{\beta^2} \right)^{n+1} \right) + c_1 \exp \left( -c_2 \delta \left( \frac{1}{\beta^2} \right)^{n+1} \right) + \exp \left( \frac{-\delta^2}{\delta + l_1 + l_2} 2^{n+1} \right) & \text{if } \beta > \frac{\sqrt{2}}{2}, \end{cases} \end{aligned} \quad (4.27)$$

where the positive constants  $c_1$  and  $c_2$  may differ term by term.

One can easily check that the coefficients of the matrix  $U_n$  are linear combinations of terms similar to  $V_{|\mathbb{T}_{n-1}|}$ , so that performing to similar calculations as before for each of them, we deduce the same deviation inequalities for  $U_n$  as in (4.27).

Now we have

$$\begin{aligned} \mathbb{P} \left( \sum_{k=0}^{n-p} \frac{1}{2^k} \left\| \sum_{C \in \{A; B\}^k} C \frac{U_{n-k}}{2^{n-k+1}} C^t \right\| > \delta \right) & \leq \mathbb{P} \left( \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A; B\}^k} \frac{1}{2^{n-k+1}} \|CU_{n-k}C^t\| > \delta \right) \\ & \leq \mathbb{P} \left( \sum_{k=p}^n \beta^{2(n-k)} \frac{1}{|\mathbb{T}_k| + 1} \|U_k\| > \delta \right) \\ & \leq \sum_{k=p}^n \mathbb{P} \left( \frac{\|U_k\|}{|\mathbb{T}_k| + 1} > \frac{\delta}{(n-p+1)\beta^{2(n-k)}} \right). \end{aligned}$$

From (4.27), we infer the following

$$\mathbb{P} \left( \sum_{k=0}^{n-p} \frac{1}{2^k} \left\| \sum_{C \in \{A; B\}^k} C \frac{U_{n-k}}{2^{n-k+1}} C^t \right\| > \delta \right) \\ \leq \begin{cases} c_1 \sum_{k=p}^n \exp \left( -c_2 \frac{\delta^2 (2\beta^4)^{k+1}}{n^2 \beta^{4n}} \right) + c_1 \sum_{k=p}^n \exp \left( -c_2 \frac{\delta (2\beta^2)^{k+1}}{n \beta^{2n}} \right) \\ \quad + c_1 \sum_{k=p}^n \exp \left( -c_2 \frac{\delta^2 2^{k+1}}{(\delta + nl \beta^{2(n-k-1)}) n \beta^{2(n-k-1)}} \right) \quad \text{if } \beta < \frac{\sqrt{2}}{2}, \\ c_1 \sum_{k=p}^n \exp \left( -c_2 \frac{\delta^2 4^n}{n^2 (k+1) 2^{k+1}} \right) + c_1 \sum_{k=p}^n \exp \left( -c_2 \frac{\delta 2^n}{(k+1)n} \right) \\ \quad + c_1 \sum_{k=p}^n \exp \left( -c_2 \frac{\delta^2 2^{k+1}}{(\delta + nl 2^{-(n-k-1)}) n 2^{-(n-k-1)}} \right) \quad \text{if } \beta = \frac{\sqrt{2}}{2}, \\ c_1 \sum_{k=p}^n \exp \left( -c_2 \frac{\delta^2 (2\beta^2)^{k+1}}{n^2 \beta^{4n}} \right) + c_1 \sum_{k=p}^n \exp \left( -c_2 \frac{\delta}{n \beta^{2n}} \right) \\ \quad + c_1 \sum_{k=p}^n \exp \left( -c_2 \frac{\delta^2 2^{k+1}}{(\delta + nl \beta^{2(n-k-1)}) n \beta^{2(n-k-1)}} \right) \quad \text{if } \beta > \frac{\sqrt{2}}{2}, \end{cases}$$

where  $l = l_1 + l_2$  and the positive constants  $c_1$  and  $c_2$  may differ term by term.

Now

- If  $\beta < \frac{\sqrt{2}}{2}$ , then on the one hand,

$$\begin{aligned} \sum_{k=p}^n \exp \left( -c \frac{\delta^2 (2\beta^4)^{k+1}}{n^2 \beta^{4n}} \right) \\ = \exp \left( -c \delta^2 \beta^4 \frac{2^{n+1}}{n^2} \right) \left( 1 + \sum_{k=p}^{n-1} \left( \exp \left( \frac{-c \delta^2}{n^2} \right) \right)^{(2\beta^4)^{k+1} \beta^{-4n} (1 - (2\beta^4)^{n-k})} \right) \\ \leq \exp \left( -c \delta^2 \beta^4 \frac{2^{n+1}}{n^2} \right) (1 + o(1)), \end{aligned}$$

where the last inequality follows from the fact that for some positive constant  $c_1$ ,

$$(2\beta^4)^{k+1} \beta^{-4n} (1 - (2\beta^4)^{n-k}) \propto c_1 (2\beta^4)^{k+1} \beta^{-4n}.$$

On the other hand, following the same lines as before, we obtain

$$\begin{aligned} \sum_{k=p}^n \exp \left( -\frac{\delta^2 2^{k+1}}{(\delta + ln \beta^{2(n-k-1)}) n \beta^{2(n-k-1)}} \right) &\leq \sum_{k=p}^n \exp \left( -c \delta^2 \frac{2^{k+1}}{n^2 \beta^{2(n-k-1)}} \right) \\ &\leq \exp \left( -c \frac{\delta^2 2^{n+1}}{(\delta + l)n^2} \right) (1 + o(1)), \end{aligned}$$

and

$$\begin{aligned} \sum_{k=p}^n \exp\left(-c \frac{\delta(2\beta^2)^{k+1}}{n\beta^{2n}}\right) &\leq \sum_{k=p}^n \exp\left(-c \frac{\delta(2\beta^2)^{k+1}}{n^2\beta^{2n}}\right) \\ &\leq \exp\left(-c\delta \frac{2^{n+1}}{n^2}\right) (1 + o(1)). \end{aligned}$$

We thus deduce that

$$\mathbb{P}\left(\sum_{k=0}^{n-p} \frac{1}{2^k} \left\| \sum_{C \in \{A;B\}^k} C \frac{U_{n-k}}{2^{n-k+1}} C^t \right\| > \delta\right) \leq c_1 \exp\left(-c_2 \delta^2 \frac{2^{n+1}}{n^2}\right) + c_1 \exp\left(-c_2 \delta \frac{2^{n+1}}{n^2}\right), \quad (4.28)$$

for some positive constants  $c_1$  and  $c_2$ .

- If  $\beta = \frac{\sqrt{2}}{2}$ , then following the same lines as before, we show that

$$\begin{aligned} \sum_{k=p}^n \exp\left(-c\delta^2 \frac{4^n}{n^2(k+1)2^{k+1}}\right) &\leq \exp\left(-c\delta^2 \frac{2^{n+1}}{n^3}\right) (1 + o(1)), \\ \sum_{k=p}^n \exp\left(-\frac{\delta^2 2^{k+1}}{(\delta + l n 2^{-(n-k-1)}) n 2^{-(n-k-1)}}\right) &\leq \exp\left(-c \frac{\delta^2 2^{n+1}}{n^2(\delta + l)}\right) (1 + o(1)), \\ \sum_{k=p}^n \exp\left(-c\delta \frac{2^n}{n(k+1)}\right) &\leq \exp\left(-c\delta \frac{2^{n+1}}{n^3}\right) (1 + o(1)). \end{aligned}$$

It then follows that

$$\mathbb{P}\left(\sum_{k=0}^{n-p} \frac{1}{2^k} \left\| \sum_{C \in \{A;B\}^k} C \frac{U_{n-k}}{2^{n-k+1}} C^t \right\| > \delta\right) \quad (4.29)$$

$$\leq c_1 \exp\left(-c_2 \delta^2 \frac{2^{n+1}}{n^3}\right) + c_1 \exp\left(-c_2 \frac{\delta^2 2^{n+1}}{n^2(\delta + l)}\right) + c_1 \exp\left(-c_2 \delta \frac{2^{n+1}}{n^3}\right), \quad (4.30)$$

for some positive constants  $c_1$  and  $c_2$ .

- If  $\beta > \frac{\sqrt{2}}{2}$ , once again following the previous lines, we get

$$\begin{aligned} \mathbb{P}\left(\sum_{k=0}^{n-p} \frac{1}{2^k} \left\| \sum_{C \in \{A;B\}^k} C \frac{U_{n-k}}{2^{n-k+1}} C^t \right\| > \delta\right) \\ \leq c_1 \exp\left(-c_2 \delta^2 \frac{1}{n^2 \beta^{2n}}\right) + c_1 \exp\left(-c_2 \frac{\delta^2}{(\delta + l) n^2 \beta^{2n}}\right) + c_1 n \exp\left(-c_2 \frac{\delta}{n^2 \beta^{2n}}\right) \end{aligned} \quad (4.31)$$

for some positive constants  $c_1$  and  $c_2$ .

We infer from the inequalities (4.28), (4.29) and (4.31) that

$$\sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C \frac{U_{n-k}}{2^{n-k+1}} C^t \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} 0.$$

□

This achieves the proof of the Proposition 4.4.

□



We now, explain the modification in the last proofs in the case 1.

**Proposition 4.10.** *Within the framework 1, we have the same conclusions as the Proposition 4.3 and 4.4 with the sequence  $(b_n)$  which satisfies condition **(V1)**.*

*Proof.* The proof follows exactly the same lines as the proof of Proposition 4.3 and 4.4, and uses the fact that if a superexponential convergence holds with a sequence  $(b_n)$  which satisfies condition **(V2)**, then it also holds with a sequence  $(b_n)$  which satisfies condition **(V1)**. We thus obtain the first convergence of (4.3), the convergences (4.8), (4.15), (4.16) and (4.12) within the framework 1 with  $(b_n)$  which satisfies condition **(V1)**. Next, following the same approach as which used to obtain (4.10), we get

$$\mathbb{P} \left( \frac{1}{|\mathbb{T}_k| + 1} \left| \sum_{i \in \mathbb{T}_{k,p}} (\varepsilon_i^2 - \sigma^2) \right| > \delta \right) \leq \begin{cases} c_1 \exp(-c_2 \delta^2 |\mathbb{T}_k|) & \text{if } \delta \text{ is small enough} \\ c_1 \exp(-c_2 \delta |\mathbb{T}_k|) & \text{if } \delta \text{ is large enough,} \end{cases} \quad (4.32)$$

where  $c_1$  and  $c_2$  are positive constants which do not depend on  $\delta$ . The first inequality holds for example if  $\delta/\gamma < \varepsilon$  and the second holds for example if  $\delta/\gamma > \varepsilon$ . On the other hand, for  $n$  large enough, let  $n_0$  such that for all  $k < n_0$ ,  $n\beta^{2(n-k)}$  is small enough so that  $\delta/(n-p+1)\gamma\beta^{2(n-k)} > \varepsilon$ . We have

$$\begin{aligned} & \mathbb{P} \left( \left\| \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C \frac{L_{n-k} - \sigma^2}{2^{n-k}} e_1 e_1^t C^t \right\| > \delta \right) \\ & \leq \sum_{k=p}^{n_0-1} \mathbb{P} \left( \frac{|L_k - \sigma^2|}{|\mathbb{T}_k| + 1} > \frac{\delta}{(n-p+1)\beta^{2(n-k)}} \right) + \sum_{k=n_0}^n \mathbb{P} \left( \frac{|L_k - \sigma^2|}{|\mathbb{T}_k| + 1} > \frac{\delta}{(n-p+1)\beta^{2(n-k)}} \right). \end{aligned}$$

Now, using (4.32) with  $\delta/(n-p+1)\beta^{2(n-k)}$  instead of  $\delta$  and following the same approach used to obtain (4.28)-(4.31) in the two sums of the right hand side of the above inequality, we are led to

$$\begin{aligned} & \mathbb{P} \left( \left\| \sum_{k=0}^{n-p} \frac{1}{2^k} \sum_{C \in \{A;B\}^k} C \frac{L_{n-k} - \sigma^2}{2^{n-k}} e_1 e_1^t C^t \right\| > \delta \right) \\ & \leq \begin{cases} c_1 \exp\left(-\frac{c_2 \delta^2 2^{n+1}}{n^2}\right) + c_1 \exp\left(-\frac{c_2 \delta 2^{n+1}}{n}\right) & \text{if } \beta \leq \frac{1}{2} \\ c_1 n \exp\left(-\frac{c_2 \delta^2}{n^2 \beta^{4n}}\right) + c_1 \exp\left(-\frac{c_2 \delta}{n \beta^{2n}}\right) & \text{if } \beta > \frac{1}{2}, \end{cases} \end{aligned}$$

and we thus obtain convergence (4.9) with  $(b_n)$  which satisfies condition **(V1)**. In the same way we obtain

$$\mathbb{P} (\|T_n^{(3)}\| > \delta) \leq \begin{cases} c_1 \exp\left(-\frac{c_2 \delta^2 2^{n+1}}{n^2}\right) + c_1 \exp\left(-\frac{c_2 \delta 2^{n+1}}{n}\right) & \text{if } \beta < \frac{1}{2}, \\ c_1 n \exp\left(-\frac{c_2 \delta^2}{n^2}\right) & \text{if } \beta = \frac{1}{2}, \\ c_1 \exp\left(-\frac{c_2 \delta^2}{n^2 \beta^{2n}}\right) + c_1 \exp\left(-\frac{c_2 \delta}{n \beta^n}\right) & \text{if } \beta > \frac{1}{2}, \end{cases}$$

so that (4.14) and then (4.13) hold for  $(b_n)$  which satisfies condition **(V1)**. To reach the convergence (4.17) and the second convergence of (4.3) with  $(b_n)$  which satisfies condition **(V1)**, we follow the same procedure as before and the proof of proposition is then complete.  $\square$

**Remark 4.11.** *Let us note that we can actually prove that*

$$\frac{1}{n} \sum_{k=2^p}^n \mathbb{X}_k \xrightarrow[b_n^2]{\text{superexp}} L_{1,2} \quad \text{and} \quad \frac{1}{n} \sum_{k=2^p}^n \mathbb{X}_k \mathbb{X}_k^t \xrightarrow[b_n^2]{\text{superexp}} L_{2,2}.$$

Indeed, let  $H_n = \sum_{k=2^{p-1}}^n \mathbb{X}_k$  and  $P_l^{(n)} = \sum_{k=2^{r_n-l}}^{\lfloor \frac{n}{2^l} \rfloor} \varepsilon_k$ . We have the following decomposition

$$\frac{H_n}{n} = \frac{1}{n} \sum_{k \in \mathbb{T}_{r_n-1, p-1}} \mathbb{X}_k + \frac{1}{n} \sum_{k=2^{r_n}}^n \mathbb{X}_k.$$

On the one hand, from Proposition 4.3, we infer that

$$\frac{1}{n} \sum_{k \in \mathbb{T}_{r_n-1, p-1}} \mathbb{X}_k \xrightarrow[b_n^2]{\text{superexp}} cL_{1,2},$$

where  $c = \lim_{n \rightarrow \infty} \frac{2^{r_n}-1}{n}$ .

On the other hand, from (2.2) we deduce that

$$\begin{aligned} \sum_{k=2^{r_n}}^n \mathbb{X}_k &= 2^{r_n-p+1} (\overline{A})^{r_n-p+1} \sum_{k=2^{p-1}}^{\lfloor \frac{n}{2^{r_n-p+1}} \rfloor} \mathbb{X}_k + 2\overline{a} \sum_{k=0}^{r_n-p} \left( \left\lfloor \frac{n}{2^k} \right\rfloor - 2^{r_n-k} + 1 \right) 2^k (\overline{A})^k e_1 \\ &\quad + \sum_{k=0}^{r_n-p} 2^k (\overline{A})^k P_k^{(n)} e_1 - \sum_{k=1}^{r_n-p+1} s_k 2^{k-1} (\overline{A})^{k-1} \left( B\mathbb{X}_{\lfloor \frac{n}{2^k} \rfloor} + \eta_{\lfloor \frac{n}{2^{k-1}} \rfloor + 1} \right), \end{aligned}$$

where

$$s_k = \begin{cases} 1 & \text{if } \lfloor \frac{n}{2^{k-1}} \rfloor \text{ is even} \\ 0 & \text{if } \lfloor \frac{n}{2^{k-1}} \rfloor \text{ is odd.} \end{cases}$$

Performing now as in the proof of Proposition 4.3, tedious but straightforward calculations lead us to

$$\frac{1}{n} \sum_{k=2^{r_n}}^n \mathbb{X}_k \xrightarrow[b_n^2]{\text{superexp}} (1-c)L_{1,2}$$

and it then follows that

$$\frac{1}{n} \sum_{k=2^p}^n \mathbb{X}_k \xrightarrow[b_n^2]{\text{superexp}} L_{1,2}.$$

The term  $\frac{1}{n} \sum_{k=2^p}^n \mathbb{X}_k \mathbb{X}_k^t$  can be dealt with in the same way.

The rest of the paper is dedicated to the proof of our main results. We focus on the proof in the case 2, and some explanation are given on how to obtain the results in the case 1.

## 5. PROOF OF THE MAIN RESULTS

We start with the proof of the deviation inequalities.

**5.1. Proof of Theorem 3.1.** We begin the proof with the case 2. Let  $\delta > 0$  and  $b > 0$  such that  $b < \|\Sigma\|/(1 + \delta)$ . We have from (2.14)

$$\begin{aligned} \mathbb{P}\left(\|\hat{\theta}_n - \theta\| > \delta\right) &= \mathbb{P}\left(\frac{\|M_n\|}{\|\Sigma_{n-1}\|} > \delta, \frac{\|\Sigma_{n-1}\|}{|\mathbb{T}_{n-1}|} \geq b\right) + \mathbb{P}\left(\frac{\|M_n\|}{\|\Sigma_{n-1}\|} > \delta, \frac{\|\Sigma_{n-1}\|}{|\mathbb{T}_{n-1}|} < b\right) \\ &\leq \mathbb{P}\left(\frac{\|M_n\|}{|\mathbb{T}_{n-1}|} > \delta b\right) + \mathbb{P}\left(\left\|\frac{\Sigma_{n-1}}{|\mathbb{T}_{n-1}|} - \Sigma\right\| > \|\Sigma\| - b\right). \end{aligned}$$

Since  $b < \|\Sigma\|/(1 + \delta)$ , then,

$$\mathbb{P}\left(\left\|\frac{\Sigma_{n-1}}{|\mathbb{T}_{n-1}|} - \Sigma\right\| > \|\Sigma\| - b\right) \leq \mathbb{P}\left(\left\|\frac{\Sigma_{n-1}}{|\mathbb{T}_{n-1}|} - \Sigma\right\| > \delta b\right).$$

It then follows that

$$\mathbb{P}\left(\|\hat{\theta}_n - \theta\| > \delta\right) \leq 2 \max\left\{\mathbb{P}\left(\frac{\|M_n\|}{|\mathbb{T}_{n-1}|} > \delta b\right), \mathbb{P}\left(\left\|\frac{\Sigma_{n-1}}{|\mathbb{T}_{n-1}|} - \Sigma\right\| > \delta b\right)\right\}.$$

On the one hand, we have

$$\begin{aligned} \mathbb{P}\left(\frac{\|M_n\|}{|\mathbb{T}_{n-1}|} > \delta b\right) &\leq \sum_{\eta=0}^1 \left\{ \mathbb{P}\left(\left|\frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} \varepsilon_{2k+\eta}\right| > \frac{\delta b}{4}\right) \right. \\ &\quad \left. + \mathbb{P}\left(\left\|\frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} \varepsilon_{2k+\eta} \mathbb{X}_k\right\| > \frac{\delta b}{4}\right) \right\}. \end{aligned}$$

Now, by carrying out the same calculations as those which have permit us to obtain Lemma 4.7 and equation (4.27), we are led to

$$\mathbb{P}\left(\frac{\|M_n\|}{|\mathbb{T}_{n-1}|} > \delta b\right) \leq \begin{cases} c_1 \exp\left(-\frac{c_2(\delta b)^2}{c_3+c_4(\delta b)} 2^n\right) & \text{if } \beta < \frac{\sqrt{2}}{2}, \\ c_1 \exp\left(-\frac{c_2(\delta b)^2}{c_3+c_4(\delta b)} \frac{2^n}{n}\right) & \text{if } \beta = \frac{\sqrt{2}}{2}, \\ c_1 \exp\left(-\frac{c_2(\delta b)^2}{c_3+c_4(\delta b)} \left(\frac{1}{\beta^2}\right)^n\right) & \text{if } \beta > \frac{\sqrt{2}}{2}, \end{cases} \quad (5.1)$$

where positive constants  $c_1, c_2, c_3$  and  $c_4$  depend on  $\sigma, \beta, \gamma$  and  $\phi$  and  $(c_3, c_4) \neq (0, 0)$ .

On the other hand, noticing that  $\Sigma_{n-1} = I_2 \otimes S_{n-1}$ , we have

$$\mathbb{P}\left(\left\|\frac{\Sigma_{n-1}}{|\mathbb{T}_{n-1}|} - \Sigma\right\| > \delta b\right) \leq 2\mathbb{P}\left(\left\|\frac{S_{n-1}}{|\mathbb{T}_{n-1}|} - L\right\| > \frac{\delta b}{2}\right).$$

Next, from the proofs of Propositions 4.3 and 4.4, we deduce that

$$\mathbb{P}\left(\left\|\frac{\Sigma_{n-1}}{|\mathbb{T}_{n-1}|} - \Sigma\right\| > \frac{b}{2}\right) \leq \begin{cases} c_1 \exp\left(-\frac{c_2(\delta b)^2}{c_3+c_4(\delta b)} \frac{2^n}{(n-1)^2}\right) & \text{if } \beta < \frac{\sqrt{2}}{2} \\ c_1 \exp\left(-\frac{c_2(\delta b)^2}{c_3+c_4(\delta b)} \frac{2^n}{(n-1)^3}\right) & \text{if } \beta = \frac{\sqrt{2}}{2} \\ c_1 \exp\left(-\frac{c_2(\delta b)^2}{c_3+c_4(\delta b)} \left(\frac{1}{(n-1)^2\beta^{2n}}\right)\right) & \text{if } \beta > \frac{\sqrt{2}}{2}, \end{cases} \quad (5.2)$$

where positive constants  $c_1, c_2, c_3$  and  $c_4$  depend on  $\sigma, \beta, \gamma$  and  $\phi$  and  $(c_3, c_4) \neq (0, 0)$ . Now, (3.1) follows from (5.1) and (5.2).

In the case 1, the proof follows exactly the same lines as before and uses the same ideas as the proof of Proposition 4.10. Particularly, we have in this case

$$\mathbb{P} \left( \left\| \frac{\Sigma_{n-1}}{|\mathbb{T}_{n-1}|} - \Sigma \right\| > \frac{b}{2} \right) \leq \begin{cases} c_1 \exp \left( -\frac{c_2(\delta b)^2}{c_3 + (\delta b)} \frac{2^n}{(n-1)^2} \right) & \text{if } \beta < \frac{1}{2} \\ c_1(n-1) \exp \left( -\frac{c_2(\delta b)^2}{c_3 + (\delta b)} \frac{2^n}{(n-1)^2} \right) & \text{if } \beta = \frac{1}{2} \\ c_1(n-1) \exp \left( -\frac{c_2(\delta b)^2}{c_3 + (\delta b)} \left( \frac{1}{(n-1)\beta^n} \right) \right) & \text{if } \beta > \frac{1}{2}. \end{cases}$$

where positive constants  $c_1, c_2$  and  $c_3$  depend on  $\sigma, \beta, \gamma$  and  $\phi$ . (3.1) then follows in this case, and this ends the proof of Theorem. 3.1.

**5.2. Proof of Theorem 3.6.** At first we need to prove the following

**Theorem 5.1.** *In the case 1 or in the case 2, the sequence  $\left( M_n / \left( b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|} \right) \right)_{n \geq 1}$  satisfies the MDP on  $\mathbb{R}^{2(p+1)}$  with speed  $b_{|\mathbb{T}_{n-1}|}^2$  and rate function*

$$I_M(x) = \sup_{\lambda \in \mathbb{R}^{2(p+1)}} \{ \lambda^t x - \lambda^t (\Gamma \otimes L) \lambda \} = \frac{1}{2} x^t (\Gamma \otimes L)^{-1} x. \quad (5.3)$$

**5.2.1. Proof of Theorem 5.1.** Now, as in Bercu et al. [7], denote by  $(\mathcal{G}_n)_{n \geq 1}$  the sister pair-wise filtration, that is  $\mathcal{G}_n = \sigma\{X_1, (X_{2k}, X_{2k+1}), 1 \leq k \leq n\}$ . We introduce the following  $(\mathcal{G}_n)$  martingale difference sequence  $(D_n)$ , given by

$$D_n = V_n \otimes Y_n = \begin{pmatrix} \varepsilon_{2n} \\ \varepsilon_{2n} \mathbb{X}_n \\ \varepsilon_{2n+1} \\ \varepsilon_{2n+1} \mathbb{X}_n \end{pmatrix}.$$

We clearly have

$$D_n D_n^t = V_n V_n^t \otimes Y_n Y_n^t.$$

So we obtain that the quadratic variation of the  $(\mathcal{G}_n)$  martingale  $(N_n)_{n \geq 2^{p-1}}$  given by

$$N_n = \sum_{k=2^{p-1}}^n D_k$$

is

$$\langle N \rangle_n = \sum_{k=2^{p-1}}^n \mathbb{E}(D_k D_k^t / \mathcal{G}_{k-1}) = \Gamma \otimes \sum_{k=2^{p-1}}^n Y_k Y_k^t.$$

Now we clearly have  $M_n = N_{|\mathbb{T}_{n-1}|}$  and  $\langle M \rangle_n = \langle N \rangle_{|\mathbb{T}_{n-1}|} = \Gamma \otimes S_{n-1}$ . From Proposition 4.1, and since  $\langle M \rangle_n = \Gamma \otimes S_{n-1}$ , we have

$$\frac{\langle M \rangle_n}{|\mathbb{T}_n|} \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} \Gamma \otimes L. \quad (5.4)$$

Before going to the proof of the MDP results, we state the exponential Lyapounov condition for  $(N_n)_{n \geq 2^{p-1}}$ , which implies exponential Lindeberg condition, that is

$$\limsup \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{k=2^{p-1}}^n \mathbb{E} \left[ \|D_k\|^2 \mathbf{1}_{\{\|D_k\| \geq r \frac{\sqrt{n}}{b_n}\}} \right] \geq \delta \right) = -\infty,$$

(see e.g [29] for more details on this implication).

**Remarks 5.2.** By [14], we infer from the condition **(Ea)** that

**(Na)** one can find  $\gamma_a > 0$  such that for all  $n \geq p-1$ , for all  $k \in \mathbb{G}_{n+1}$  and for all  $t \in \mathbb{R}$ , with  $\mu_a = \mathbb{E}(|\varepsilon_k|^a / \mathcal{F}_n)$  a.s.

$$\mathbb{E} [\exp t (|\varepsilon_k|^a - \mu_a) / \mathcal{F}_n] \leq \exp \left( \frac{\gamma_a t^2}{2} \right) \quad a.s.$$

**Proposition 5.3.** Let  $(b_n)$  a sequence satisfying the Assumption **(V2)**. Assume that hypothesis **(Na)** and **(Xa)** are satisfied. Then there exists  $B > 0$  such that

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{j=2^{p-1}}^n \mathbb{E} [\|D_j\|^a / \mathcal{G}_{j-1}] > B \right) = -\infty.$$

*Proof.* We are going to prove that

$$\limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_n|}^2} \log \mathbb{P} \left( \frac{1}{|\mathbb{T}_n|} \sum_{j=2^p}^{|\mathbb{T}_n|} \mathbb{E} [\|D_j\|^a / \mathcal{G}_{j-1}] > B \right) = -\infty, \quad (5.5)$$

and the Proposition (5.3) will follow performing as in Remark 4.11. We have

$$\sum_{j \in \mathbb{T}_{n,p}} \mathbb{E} [\|D_j\|^a / \mathcal{G}_{j-1}] \leq c \mu^a \sum_{j \in \mathbb{T}_{n,p}} (1 + \|\mathbb{X}_j\|^a),$$

where  $c$  is a positive constant which depends on  $a$ . From (2.2), we deduce that

$$\sum_{j \in \mathbb{T}_{n,p}} \|\mathbb{X}_j\|^a \leq \frac{c^2}{(1-\beta)^{a-1}} P_n + \frac{c^2 \alpha^a Q_n}{(1-\beta)^{a-1}} + 2c R_n \bar{X}_1^a,$$

where

$$P_n = \sum_{j \in \mathbb{T}_{n,p}} \sum_{i=0}^{r_j-p} \beta^i |\varepsilon_{[\frac{j}{2^i}]}|^a, \quad Q_n = \sum_{j \in \mathbb{T}_{n,p}} \sum_{i=0}^{r_j-p} \beta^i, \quad R_n = \sum_{j \in \mathbb{T}_{n,p}} \beta^{a(r_j-p+1)},$$

and  $c$  is a positive constant. Now, performing as in the proof of Proposition 4.4, using hypothesis **(Na)** and **(Xa)** instead of **(N2)** and **(X2)** we get for  $B$  large enough

$$\limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_n|}^2} \log \mathbb{P} \left( \frac{1}{|\mathbb{T}_n|} \sum_{j \in \mathbb{T}_{n,p}} \|\mathbb{X}_j\|^a > B \right) = -\infty. \quad (5.6)$$

Now (5.6) leads us to (5.5) and performing as in Remark 4.11, we obtain the Proposition 5.3.

**Remarks 5.4.** *In the case 1, we clearly have that  $(\mathbb{X}_n, n \in \mathbb{T}_{\cdot, p-1})$ , where*

$$\mathbb{T}_{\cdot, p-1} = \bigcup_{r=p-1}^{\infty} \mathbb{G}_r,$$

*is a bifurcating Markov chain with initial state  $\mathbb{X}_{2^{p-1}} = (X_{2^{p-1}}, X_{2^{p-2}}, \dots, X_1)^t$ . Let  $\nu$  the law of  $\mathbb{X}_{2^{p-1}}$ . From hypothesis **(X2)**, we deduce that  $\nu$  has finite moments of all orders. We denote by  $P$  the transition probability kernel associated to  $(\mathbb{X}_n, n \in \mathbb{T}_{\cdot, p-1})$ . Let  $(\mathbb{Y}_r, r \in \mathbb{N})$  the ergodic stable Markov chain associated to  $(\mathbb{X}_n, n \in \mathbb{T}_{\cdot, p-1})$ . This Markov chain is defined as follows, starting from the root  $\mathbb{Y}_0 = \mathbb{X}_{2^{p-1}}$  and if  $\mathbb{Y}_r = \mathbb{X}_n$  then  $\mathbb{Y}_{r+1} = \mathbb{X}_{2n+\zeta_{r+1}}$  for a sequence of independent Bernoulli r.v.  $(\zeta_q, q \in \mathbb{N}^*)$  such that  $\mathbb{P}(\zeta_q = 0) = \mathbb{P}(\zeta_q = 1) = 1/2$ . Let  $\mu$  the stationary distribution associated to  $(\mathbb{Y}_r, r \in \mathbb{N})$ . For more details on bifurcating Markov chain and the associated ergodic stable Markov chain, we refer to [19] (see also [9]).*

*From [9], we deduce that for all real bounded function  $f$  defined on  $(\mathbb{R}^p)^3$ ,*

$$\frac{1}{b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}} \sum_{k \in \mathbb{T}_{n-1, p-1}} f(\mathbb{X}_k, \mathbb{X}_{2k}, \mathbb{X}_{2k+1})$$

*satisfies a MDP on  $\mathbb{R}$  with speed  $b_{|\mathbb{T}_{n-1}|}^2$  and the rate function  $I(x) = \frac{x^2}{2S^2(f)}$ , where  $S^2(f) = < \mu, P(f^2) - (Pf)^2 >$ .*

*Now, let the function  $f$  defined on  $(\mathbb{R}^p)^3$  by  $f(x, y, z) = \|x\|^2 + \|y\|^2 + \|z\|^2$ . Then, using the relation (4.1) in Proposition 4.1, the above MDP for real bounded functionals of the bifurcating Markov chain  $(\mathbb{X}_n, n \in \mathbb{T}_{\cdot, p-1})$  and the truncation of the function  $f$ , we prove (in the same manner as the proof of lemma 3 in Worms [30]) that for all  $r > 0$*

$$\limsup_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{j=2^{p-1}}^n (\|X_j\|^2 + \|X_{2j}\|^2 + \|X_{2j+1}\|^2) \times \mathbf{1}_{\{\|\mathbb{X}_j\| + \|\mathbb{X}_{2j}\| + \|\mathbb{X}_{2j+1}\| > R\}} > r \right) = -\infty,$$

*which implies the following Lindeberg condition (for more detail one can see Proposition 2 in Worms [30])*

$$\limsup_{n \rightarrow \infty} \frac{1}{b_n^2} \log \mathbb{P} \left( \frac{1}{n} \sum_{j=2^{p-1}}^n (\|X_j\|^2 + \|X_{2j}\|^2 + \|X_{2j+1}\|^2) \times \mathbf{1}_{\{\|\mathbb{X}_j\| + \|\mathbb{X}_{2j}\| + \|\mathbb{X}_{2j+1}\| > r \frac{\sqrt{n}}{b_n}\}} > \delta \right) = -\infty,$$

*for all  $\delta > 0$  and for all  $r > 0$ . Notice that the above Lindeberg condition implies particularly the Lindeberg condition on the sequence  $(\mathbb{X}_n)$ .*

Now, we back to the proof of Theorem 5.1. We divide the proof into four steps. In the first one, we introduce a truncation of the martingale  $(M_n)_{n \geq 0}$  and prove that the truncated martingale satisfies some MDP thanks to Puhalskii's Theorem 3.11. In the second part, we show that the truncated martingale is an exponentially good approximation of  $(M_n)$ , see e.g. Definition 4.2.14 in [13]. We conclude by the identification of the rate function.

## Proof in the case 2

**Step 1.** From now on, in order to apply Puhalskii's result [22] (Puhalskii's Theorem 3.11) for the MDP for martingales, we introduce the following truncation of the martingale  $(M_n)_{n \geq 0}$ . For  $r > 0$  and  $R > 0$ ,

$$M_n^{(r,R)} = \sum_{k \in \mathbb{T}_{n-1,p-1}} D_{k,n}^{(r,R)}.$$

where, for all  $1 \leq k \leq n$ ,  $D_{k,n}^{(r,R)} = V_k^{(R)} \otimes Y_{k,n}^{(r)}$ , with

$$V_n^{(R)} = \left( \varepsilon_{2n}^{(R)}, \varepsilon_{2n+1}^{(R)} \right)^t \quad \text{and} \quad Y_{k,n}^{(r)} = \left( 1, \mathbb{X}_{k,n}^{(r)} \right)^t,$$

where

$$\varepsilon_k^{(R)} = \varepsilon_k \mathbf{1}_{\{|\varepsilon_k| \leq R\}} - \mathbb{E} \left[ \varepsilon_k \mathbf{1}_{\{|\varepsilon_k| \leq R\}} \right], \quad \mathbb{X}_{k,n}^{(r)} = \mathbb{X}_k \mathbf{1}_{\left\{ \|\mathbb{X}_k\| \leq r \frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}} \right\}}.$$

We introduce  $\Gamma^{(R)}$  the conditional covariance matrix associated with  $(\varepsilon_{2k}^{(R)}, \varepsilon_{2k+1}^{(R)})^t$  and the truncated matrix associated with  $S_n$  :

$$\Gamma^{(R)} = \begin{pmatrix} \sigma_R^2 & \rho_R \\ \rho_R & \sigma_R^2 \end{pmatrix} \quad \text{and} \quad S_n^{(r)} = \sum_{k \in \mathbb{T}_{n,p-1}} \begin{pmatrix} 1 & (\mathbb{X}_{k,n}^{(r)})^t \\ \mathbb{X}_{k,n}^{(r)} & \mathbb{X}_{k,n}^{(r)} (\mathbb{X}_{k,n}^{(r)})^t \end{pmatrix}.$$

The condition **(P2)** in Puhalskii's Theorem 3.11 is verified by the construction of the truncated martingale, that is for some positive constant  $c$ , we have that for all  $k \in \mathbb{T}_{n-1}$

$$\|D_{k,n}^{(r,R)}\| \leq c \frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}}.$$

From Proposition 5.3, we also have for all  $r > 0$ ,

$$\frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1,p-1}} \mathbb{X}_k \mathbf{1}_{\left\{ \|\mathbb{X}_k\| > r \frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}} \right\}} \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0; \quad (5.7)$$

and

$$\frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1,p-1}} \mathbb{X}_k \mathbb{X}_k^t \mathbf{1}_{\left\{ \|\mathbb{X}_k\| > r \frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}} \right\}} \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0. \quad (5.8)$$

From (5.7) and (5.8), we deduce that for all  $r > 0$

$$\frac{1}{|\mathbb{T}_{n-1}|} \left( S_{n-1} - S_{n-1}^{(r)} \right) \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0. \quad (5.9)$$

Then, we easily transfer the properties (5.4) to the truncated martingale  $(M_n^{(r,R)})_{n \geq 0}$ . We have for all  $R > 0$  and all  $r > 0$ ,

$$\frac{\langle M^{(r,R)} \rangle_n}{|\mathbb{T}_{n-1}|} = \Gamma^{(R)} \otimes \frac{S_{n-1}^{(r)}}{|\mathbb{T}_{n-1}|} = -\Gamma^{(R)} \otimes \left( \frac{S_{n-1} - S_{n-1}^{(r)}}{|\mathbb{T}_{n-1}|} \right) + \Gamma^{(R)} \otimes \frac{S_{n-1}}{|\mathbb{T}_{n-1}|} \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} \Gamma^{(R)} \otimes L$$

That is condition **(P1)** in Puhalskii's Theorem 3.11.

Note also that Proposition 5.3 work for the truncated martingale  $(M_n^{(r,R)})_{n \geq 0}$ , which ensures the Lindeberg's condition and thus condition **(P3)** to  $(M_n^{(r,R)})_{n \geq 0}$ . By Theorem 3.11 in the Appendix, we deduce that  $(M_n^{(r,R)} / (b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}))_{n \geq 0}$  satisfies a MDP on  $\mathbb{R}^{2(p+1)}$  with speed  $b_{|\mathbb{T}_{n-1}|}^2$  and good rate function given by

$$I_R(x) = \frac{1}{2}x^t(\Gamma^{(R)} \otimes L)^{-1}x. \quad (5.10)$$

**Step 2.** At first, we infer from the hypothesis **(Ea)** that:

**(N1R)** there is a sequence  $(\kappa_R)_{R>0}$  with  $\kappa_R \rightarrow 0$  when  $R$  goes to infinity, such that for all  $n \geq p-1$ , for all  $k \in \mathbb{G}_{n+1}$ , for all  $t \in \mathbb{R}$  and for  $R$  large enough

$$\mathbb{E} [\exp t (\varepsilon_k - \varepsilon_k^{(R)}) / \mathcal{F}_n] \leq \exp \left( \frac{\kappa_R t^2}{2} \right), \quad a.s.$$

The approximation, in the sense of the moderate deviation, is described by the following convergence, for all  $r > 0$  and all  $\delta > 0$ ,

$$\limsup_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( \frac{\|M_n - M_n^{(r,R)}\|}{\sqrt{|\mathbb{T}_{n-1}|} b_{|\mathbb{T}_{n-1}|}} > \delta \right) = -\infty.$$

For that, we shall prove that for  $\eta \in \{0, 1\}$

$$I_1 = \frac{1}{\sqrt{|\mathbb{T}_{n-1}|} b_{|\mathbb{T}_{n-1}|}} \sum_{k \in \mathbb{T}_{n-1, p-1}} \left( \varepsilon_{2k+\eta} - \varepsilon_{2k+\eta}^{(R)} \right) \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0, \quad (5.11)$$

$$I_2 = \frac{1}{\sqrt{|\mathbb{T}_{n-1}|} b_{|\mathbb{T}_{n-1}|}} \sum_{k \in \mathbb{T}_{n-1, p-1}} \left( \varepsilon_{2k+\eta} \mathbb{X}_k - \varepsilon_{2k+\eta}^{(R)} \mathbb{X}_{k,n}^{(r)} \right) \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0. \quad (5.12)$$

To prove (5.11) and (5.12), we have to do it only for  $\eta = 0$  the same proof works for  $\eta = 1$ .

**Proof of (5.11)** We have for all  $\alpha > 0$  and  $R$  large enough

$$\begin{aligned} & \mathbb{E} \left( \exp \left( \alpha \sum_{k \in \mathbb{T}_{n-1, p-1}} (\varepsilon_{2k} - \varepsilon_{2k}^{(R)}) \right) \right) \\ &= \mathbb{E} \left[ \prod_{k \in \mathbb{T}_{n-2, p-1}} \exp \left( \alpha (\varepsilon_{2k} - \varepsilon_{2k}^{(R)}) \right) \times \mathbb{E} \left[ \prod_{k \in \mathbb{G}_{n-1}} \exp \left( \alpha (\varepsilon_{2k} - \varepsilon_{2k}^{(R)}) \right) / \mathcal{F}_{n-1} \right] \right] \\ &= \mathbb{E} \left[ \prod_{k \in \mathbb{T}_{n-2, p-1}} \exp \left( \alpha (\varepsilon_{2k} - \varepsilon_{2k}^{(R)}) \right) \times \prod_{k \in \mathbb{G}_{n-1}} \mathbb{E} \left[ \exp \left( \alpha (\varepsilon_{2k} - \varepsilon_{2k}^{(R)}) \right) / \mathcal{F}_{n-1} \right] \right] \\ &\leq \mathbb{E} \left[ \prod_{k \in \mathbb{T}_{n-2, p-1}} \exp \left( \alpha (\varepsilon_{2k} - \varepsilon_{2k}^{(R)}) \right) \exp (|\mathbb{G}_{n-1}| \alpha^2 \kappa_R) \right] \\ &\leq \exp (|\mathbb{T}_{n-1}| \alpha^2 \kappa_R). \end{aligned}$$

where hypothesis **(N1R)** was used to get the first inequality, and the second was obtained by induction. By Chebyshev inequality and the previous calculation applied to  $\alpha = \lambda b_{|\mathbb{T}_{n-1}|} / |\mathbb{T}_{n-1}|$ , we obtain for all  $\delta > 0$

$$\mathbb{P} \left( \frac{1}{\sqrt{|\mathbb{T}_{n-1}|} b_{|\mathbb{T}_{n-1}|}} \sum_{k \in \mathbb{T}_{n-1, p-1}} (\varepsilon_{2k} - \varepsilon_{2k}^{(R)}) \geq \delta \right) \leq \exp \left( -b_{|\mathbb{T}_{n-1}|}^2 (\delta \lambda - \kappa_R \lambda^2) \right).$$



Optimizing on  $\lambda$ , we obtain

$$\frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( \frac{1}{\sqrt{|\mathbb{T}_{n-1}|} b_{|\mathbb{T}_{n-1}|}} \sum_{k \in \mathbb{T}_{n-1, p-1}} \left( \varepsilon_{2k} - \varepsilon_{2k}^{(R)} \right) \geq \delta \right) \leq -\frac{\delta^2}{4\kappa_R}.$$

Letting  $n$  goes to infinity and then  $R$  goes to infinity, we obtain the negligibility in (5.11).

**Proof of (5.12)** Now, since we have the decomposition

$$\varepsilon_{2k} \mathbb{X}_k - \varepsilon_{2k}^{(R)} \mathbb{X}_{k,n}^{(r)} = \left( \varepsilon_{2k} - \varepsilon_{2k}^{(R)} \right) \mathbb{X}_{k,n}^{(r)} + \varepsilon_{2k} \left( \mathbb{X}_k - \mathbb{X}_{k,n}^{(r)} \right),$$

we introduce the following notations

$$L_n^{(r)} = \sum_{k \in \mathbb{T}_{n-1, p-1}} \varepsilon_{2k} \left( \mathbb{X}_k - \mathbb{X}_{k,n}^{(r)} \right) \quad \text{and} \quad F_n^{(r, R)} = \sum_{k \in \mathbb{T}_{n-1, p-1}} \left( \varepsilon_{2k} - \varepsilon_{2k}^{(R)} \right) \mathbb{X}_{k,n}^{(r)}.$$

To prove (5.12), we will show that for all  $r > 0$

$$\frac{L_n^{(r)}}{\sqrt{|\mathbb{T}_{n-1}|} b_{|\mathbb{T}_{n-1}|}} \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0, \quad (5.13)$$

and for all  $r > 0$  and all  $\delta > 0$

$$\limsup_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( \frac{\|F_n^{(r, R)}\|}{b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}} > \delta \right) = -\infty. \quad (5.14)$$

Let us first deal with  $(L_n^{(r)})$ . Let its first component

$$L_{n,1}^{(r)} = \sum_{k \in \mathbb{T}_{n-1, p-1}} \varepsilon_{2k} \left( X_k - X_{k,n}^{(r)} \right).$$

For  $\lambda \in \mathbb{R}$ , we consider the random sequence  $(Z_{n,1}^{(r)})_{n \geq p-1}$  defined by

$$Z_{n,1}^{(r)} = \exp \left( \lambda L_{n,1}^{(r)} - \frac{\lambda^2 \phi}{2} \sum_{k \in \mathbb{T}_{n-1, p-1}} X_k^2 \mathbf{1}_{\left\{ \|\mathbb{X}_k\| > r \frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}} \right\}} \right)$$

where  $\phi$  appears in **(N1)**.

For  $b > 0$ , we introduce the following event

$$A_{n,1}^{(r)}(b) = \left\{ \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} X_k^2 \mathbf{1}_{\left\{ \|\mathbb{X}_k\| > r \frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}} \right\}} > b \right\}.$$

Using **(N1)**, we have for all  $\delta > 0$

$$\begin{aligned} & \mathbb{P} \left( \frac{1}{b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}} L_{n,1}^{(r)} > \delta \right) \\ & \leq \mathbb{P} \left( A_{n,1}^{(r)}(b) \right) + \mathbb{P} \left( Z_{n,1}^{(r)} > \exp \left( \delta \lambda b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|} - \frac{\lambda^2 \phi}{2} b_{|\mathbb{T}_{n-1}|} \right) \right) \\ & \leq \mathbb{P} \left( A_{n,1}^{(r)}(b) \right) + \exp \left( -b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|} \left( \delta \lambda - \frac{b \phi \sqrt{|\mathbb{T}_{n-1}|}}{2 b_{|\mathbb{T}_{n-1}|}} \lambda^2 \right) \right), \end{aligned} \quad (5.15)$$

where the second term in (5.15) is obtained by conditioning successively on  $(\mathcal{G}_i)_{2^{p-1} \leq i \leq |\mathbb{T}_{n-1}|-1}$  and using the fact that

$$\mathbb{E} \left[ \exp \left( \lambda \varepsilon_{2^p} \left( X_{2^{p-1}} - X_{2^{p-1}}^{(r)} \right) - \frac{\lambda^2 \phi}{2} X_{2^{p-1}}^2 \mathbf{1}_{\left\{ \|\mathbb{X}_{2^{p-1}}\| > r \frac{\sqrt{2^{p-1}}}{b_{2^{p-1}}} \right\}} \right) \right] \leq 1,$$

which follows from **(N1)**.

From Proposition 5.3, we have for all  $b > 0$

$$\limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( A_{n,1}^{(r)}(b) \right) = -\infty,$$

so that taking  $\lambda = \delta b_{|\mathbb{T}_{n-1}|} / (b\phi\sqrt{|\mathbb{T}_{n-1}|})$  in (5.15), we are led to

$$\limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( \frac{L_{n,1}^{(r)}}{b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}} > \delta \right) \leq -\frac{\delta^2}{2b\phi}.$$

Letting  $b \rightarrow 0$ , we obtain that the right hand of the last inequality goes to  $-\infty$ . Proceeding in the same way for  $-L_{n,1}^{(r)}$ , we deduce that for all  $r > 0$

$$\frac{L_{n,1}^{(r)}}{b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}} \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0.$$

Now, it is easy to check that the same proof works for the others components of  $L_n^{(r)}$ . We thus conclude the proof of (5.13).

Eventually, let us treat the term  $(F_n^{(r,R)})$ . We follow the same approach as in the proof of (5.13). Let its first component

$$F_{n,1}^{(r,R)} = \sum_{k \in \mathbb{T}_{n-1,p-1}} (\varepsilon_{2k} - \varepsilon_{2k}^{(R)}) X_{k,n}^{(r)}$$

For  $\lambda \in \mathbb{R}$ , we consider the random sequence  $(W_{n,1}^{(r,R)})_{n \geq p-1}$  defined by

$$W_{n,1}^{(r,R)} = \exp \left( \lambda \sum_{k \in \mathbb{T}_{n-1,p-1}} (\varepsilon_{2k} - \varepsilon_{2k}^{(R)}) X_{k,n}^{(r)} - \frac{\lambda^2 \kappa_R}{2} \sum_{k \in \mathbb{T}_{n-1,p-1}} (X_{k,n}^{(r)})^2 \right)$$

where  $\kappa_R$  appears in **(N1R)**.

Let  $b > 0$ . Consider the following event  $B_{n,1}^{(r)}(b) = \left\{ \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1,p-1}} (X_{k,n}^{(r)})^2 > b \right\}$ .

We have for all  $\delta > 0$ ,

$$\begin{aligned} & \mathbb{P} \left( \frac{F_{n,1}^{(r,R)}}{b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}} > \delta \right) \\ & \leq \mathbb{P} \left( B_{n,1}^{(r)}(b) \right) + \mathbb{P} \left( W_{n,1}^{(r,R)} > \exp \left( \delta \lambda b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|} - \frac{\lambda^2 \kappa_R}{2} |\mathbb{T}_{n-1}| b \right) \right) \\ & \leq \mathbb{P} \left( B_{n,1}^{(r)}(b) \right) + \exp \left( -b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|} \left( \delta \lambda - \frac{b \kappa_R \sqrt{|\mathbb{T}_{n-1}|}}{2 b_{|\mathbb{T}_{n-1}|}} \lambda^2 \right) \right) \end{aligned} \quad (5.16)$$

where the second term in (5.16) is obtained by conditioning successively on  $(\mathcal{G}_i)_{2^{p-1} \leq i \leq |\mathbb{T}_{n-1}|-1}$  and using the fact that

$$\mathbb{E} \left[ \exp \left( \lambda \left( \varepsilon_{2^p} - \varepsilon_{2^p}^{(R)} \right) X_{2^{p-1}}^{(r)} - \frac{\lambda^2 \kappa_R}{2} \left( X_{2^{p-1}}^{(r)} \right)^2 \right) \right] \leq 1,$$

Since  $B_{n,1}^{(r)}(b) \subset \left\{ \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1,p-1}} X_k^2 > b \right\}$ , from Proposition 4.4, we deduce that for  $b$  large enough

$$\limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( B_{n,1}^{(r)}(b) \right) = -\infty,$$

so that choosing  $\lambda = \delta b_{|\mathbb{T}_{n-1}|} / (\kappa_R b \sqrt{|\mathbb{T}_{n-1}|})$ , we get for all  $\delta > 0$

$$\limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( \frac{F_{n,1}^{(r,R)}}{b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}} > \delta \right) \leq -\frac{\delta^2}{2\kappa_R b}.$$

Letting  $R$  to infinity, we obtain that

$$\limsup_{R \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( \frac{F_{n,1}^{(r,R)}}{b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}} > \delta \right) = -\infty.$$

Now it is easy to check that the same works for  $-F_{n,1}^{(r,R)}$  and for the others components of  $F_n^{(r,R)}$ . We thus conclude (5.14) for all  $r > 0$ .

**Step 3.** By application of Theorem 4.2.16 in [13], we find that  $(M_n / (b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}))$  satisfies an MDP on  $\mathbb{R}^{2(p+1)}$  with speed  $b_{|\mathbb{T}_{n-1}|}^2$  and rate function

$$\tilde{I}(x) = \sup_{\delta > 0} \liminf_{R \rightarrow \infty} \inf_{z \in B_{x,\delta}} I_R(z),$$

where  $I_R$  is given in (5.10) and  $B_{x,\delta}$  denotes the ball  $\{z : |z - x| < \delta\}$ . The identification of the rate function  $\tilde{I} = I_M$ , where  $I_M$  is given in (5.3) is done easily (see for example [16]), which concludes the proof of Theorem 5.1.

### Proof in the case 1.

For the proof in the case 1, there are no change in Step 1, and Step 3, instead of (5.7), (5.8), and **(N1)**, we use Remark 5.4 and **(G1)**. In Step 2, the negligibility in (5.11), comes from the MDP of the i.i.d. sequences  $(\varepsilon_{2k} - \varepsilon_{2k}^{(R)})$  since it verifies the condition, for  $\lambda > 0$  and all  $R > 0$

$$\mathbb{E}(\exp(\lambda(\varepsilon_{2k} - \varepsilon_{2k}^{(R)}))) < \infty.$$

The negligibility of  $(L_n^{(r)})$  works in the same way. For  $(F_n^{(r,R)})$  we will use the MDP for martingale, see Proposition 3.10. For  $R$  large enough, we have

$$\begin{aligned} \mathbb{P} \left( \left| X_{k,n}^{(r)} \left( \varepsilon_{2k} - \varepsilon_{2k}^{(R)} \right) \right| > b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|} \mid \mathcal{F}_{k-1} \right) &\leq \mathbb{P} \left( \left| \varepsilon_{2k} - \varepsilon_{2k}^{(R)} \right| > \frac{b_{|\mathbb{T}_{n-1}|}^2}{r} \right), \\ &= \mathbb{P} \left( \left| \varepsilon_2 - \varepsilon_2^{(R)} \right| > \frac{b_{|\mathbb{T}_{n-1}|}^2}{r} \right) = 0. \end{aligned}$$

This implies that

$$\limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \left( |\mathbb{T}_{n-1}| \operatorname{ess\,sup}_{k \geq 1} \mathbb{P} \left( \left| X_{k,n}^{(r)} \left( \varepsilon_{2k} - \varepsilon_{2k}^{(R)} \right) \right| > b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|} \mid \mathcal{F}_{k-1} \right) \right) = -\infty.$$

That is condition **(D2)** in Proposition 3.10.

For all  $\gamma > 0$  and all  $\delta > 0$ , we obtain from Remark 5.4, that

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} \left( X_{k,n}^{(r)} \right)^2 \mathbf{I}_{\left\{ |X_{k,n}^{(r)}| > \gamma \frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}} \right\}} > \delta \right) \\ & \leq \limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} X_k^2 \mathbf{1}_{\left\{ |X_k| > \gamma \frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}} \right\}} > \delta \right) = -\infty. \end{aligned}$$

That is condition **(D3)** in Proposition 3.10. Finally, from Remark 5.4 and in the same way as in (5.9), it follows that

$$\frac{\langle F^{(r,R)} \rangle_{n,1}}{|\mathbb{T}_{n-1}|} = Q_R \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} (X_{k,n}^{(r)})^2 \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} Q_R \ell$$

for some positive constant  $\ell$ , where  $Q_R = \mathbb{E} \left[ \left( \varepsilon_2 - \varepsilon_2^{(R)} \right)^2 \right]$ . That is condition **(D1)** in Proposition 3.10. Moreover, it is clear that  $Q_R$  converges to 0 as  $R$  goes to infinity. In light of foregoing, we infer from Proposition 3.10, that  $(F_{n,1}^{(r,R)} / (b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}))$  satisfies an MDP on  $\mathbb{R}$  of speed  $b_{|\mathbb{T}_{n-1}|}^2$  and rate function  $I_R(x) = x^2 / (2Q_R \ell)$ . In particular, this implies that for all  $\delta > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{b_{|\mathbb{T}_{n-1}|}^2} \log \mathbb{P} \left( \frac{|F_{n,1}^{(r,R)}|}{b_{|\mathbb{T}_{n-1}|} \sqrt{|\mathbb{T}_{n-1}|}} > \delta \right) \leq -\frac{\delta^2}{2Q_R \ell},$$

and letting  $R$  go to infinity clearly leads to the result.

**5.2.2. Proof of Theorem 3.4.** The proof works in the case 1 and in the case 2. From (2.14), we have

$$\frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}} (\hat{\theta}_n - \theta) = |\mathbb{T}_{n-1}| \Sigma_{n-1}^{-1} \frac{M_n}{b_{|\mathbb{T}_{n-1}|} |\mathbb{T}_{n-1}|}$$

From Proposition 4.1, we obtain that

$$\frac{\Sigma_n}{|\mathbb{T}_n|} = I_2 \otimes \frac{S_n}{|\mathbb{T}_n|} \xrightarrow[b_{|\mathbb{T}_n|}^2]{\text{superexp}} I_2 \otimes L. \quad (5.17)$$

According to Lemma 4.1 of [29], together with (5.17), we deduce that

$$|\mathbb{T}_{n-1}| \Sigma_{n-1}^{-1} \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} I_2 \otimes L^{-1}. \quad (5.18)$$

From Theorem 5.1, (5.18) and the contraction principle [13], we deduce that the sequence  $(\sqrt{|\mathbb{T}_{n-1}|} (\hat{\theta}_n - \theta) / b_{|\mathbb{T}_{n-1}|})_{n \geq 1}$  satisfies the MDP with rate function  $I_\theta$  given by (3.3).  $\square$

### 5.3. Proof of Theorem 3.6.

For the proof of Theorem 3.6, the case 1 is an easy consequence of the classical MDP for i.i.d.r.v. applied to the sequence  $(\varepsilon_{2k}^2 + \varepsilon_{2k+1}^2)$ , for the case 2, we will use Proposition 3.10, rather than Puhalskii's Theorem 3.11.

We will prove that the sequence  $\left(\sqrt{|\mathbb{T}_{n-1}|}(\sigma_n^2 - \sigma^2)/b_{|\mathbb{T}_{n-1}|}\right)$  satisfies the MDP. For that, we will prove that conditions **(D1)**, **(D2)** and **(D3)** of Proposition 3.10 are verified. Let us consider the  $\mathcal{G}_n$ -martingale  $(N_n)_{n \geq 2^{p-1}}$  given by

$$N_n = \sum_{k=2^{p-1}}^n \nu_k, \quad \text{where } \nu_k = \varepsilon_{2k}^2 + \varepsilon_{2k+1}^2 - 2\sigma^2.$$

It is easy to see that its predictable quadratic variation is given by

$$\langle N \rangle_n = \sum_{k=2^{p-1}}^n \mathbb{E} [\nu_k^2 / \mathcal{G}_{k-1}] = (n - 2^{p-1} + 1)(2\tau^4 - 4\sigma^4 + 2\nu^2),$$

which immediately implies that

$$\frac{\langle N \rangle_n}{n} \xrightarrow[b_n^2]{\text{superexp}} 2\tau^4 - 4\sigma^4 + 2\nu^2,$$

ensuring condition **(D1)** in Proposition 3.10.

Next, for  $B > 0$  large enough, we have for  $a > 2$  (in **(Ea)**), and some positive constant  $c$

$$\mathbb{P} \left( \frac{1}{n} \sum_{k=2^{p-1}}^n |\nu_k|^a > B \right) \leq 3 \max_{\eta \in \{0,1\}} \left\{ \mathbb{P} \left( \frac{1}{n} \sum_{k=2^{p-1}}^n |\varepsilon_{2k+\eta}|^{2a} > \frac{B}{3c} \right) \right\}.$$

From hypothesis **(Ea)** and since  $B$  is large enough, we obtain, for a suitable  $t > 0$  via the Chernoff inequality and several successive conditioning on  $(\mathcal{G}_n)$ , for  $\eta \in \{0,1\}$

$$\mathbb{P} \left( \frac{1}{n} \sum_{k=2^{p-1}}^n |\varepsilon_{2k+\eta}|^{2a} > \frac{B}{3c} \right) \leq \exp \left( -tn \left( \frac{B}{3c} - \log E \right) \right) \leq \exp(-tc'n),$$

where  $c, c'$  are a positive generic constant. Therefore, for  $B > 0$  large enough, we deduce that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \left( \frac{1}{n} \sum_{k=2^{p-1}}^n |\nu_k|^a > B \right) < 0,$$

and this implies (see e.g [29]) exponential Lindeberg condition, that is for all  $r > 0$

$$\frac{1}{n} \sum_{k=2^{p-1}}^n \nu_k^2 \mathbf{1}_{\{|\nu_k| > r \frac{\sqrt{n}}{b_n}\}} \xrightarrow[b_n^2]{\text{superexp}} 0.$$

That is condition **(D3)** in Proposition 3.10.

Now, for all  $k \in \mathbb{N}$  and a suitable  $t > 0$  we have

$$\begin{aligned} \mathbb{P}(|\nu_k| > b_n \sqrt{n}/\mathcal{G}_{k-1}) &\leq \sum_{\eta=0}^1 \mathbb{P}\left(|\varepsilon_{2k+\eta}^2 - \sigma^2| > \frac{b_n \sqrt{n}}{2}/\mathcal{G}_{k-1}\right) \\ &\leq \exp\left(\frac{-tb_n \sqrt{n}}{2}\right) \sum_{\eta=0}^1 \mathbb{E}\left[\exp(t|\varepsilon_{2k+\eta}^2 - \sigma^2|)/\mathcal{G}_{k-1}\right] \\ &\leq 2E' \exp\left(\frac{-tb_n \sqrt{n}}{2}\right), \end{aligned}$$

where from hypothesis **(Na)**,  $E'$  is finite and positive. We are thus led to

$$\frac{1}{b_n^2} \log \left( n \operatorname{ess\,sup}_{k \in \mathbb{N}^*} \mathbb{P}(|\nu_k| > b_n \sqrt{n}/\mathcal{G}_{k-1}) \right) \leq \frac{\log(2E'n)}{b_n^2} - \frac{t\sqrt{n}}{b_n},$$

and consequently, letting  $n$  goes to infinity, we get the condition **(D2)** in Proposition 3.10.

Now, applying Proposition 3.10, we conclude that  $(N_n/(b_n \sqrt{n}))_{n \geq 0}$  satisfies the MDP with speed  $b_n^2$  and rate function

$$I_N(x) = \frac{x^2}{4(\tau^4 - 2\sigma^4 + 2\nu^2)}.$$

Applying the foregoing to  $|\mathbb{T}_{n-1}|$  and using contraction principle (see e.g [13]), we deduce that the sequence

$$\frac{\sqrt{|\mathbb{T}_{n-1}|}}{b_{|\mathbb{T}_{n-1}|}}(\sigma_n^2 - \sigma^2) = \frac{N_{|\mathbb{T}_{n-1}|}}{2b_{|\mathbb{T}_{n-1}|}\sqrt{|\mathbb{T}_{n-1}|}}$$

satisfies a MDP with speed  $b_{|\mathbb{T}_{n-1}|}^2$  and rate function  $I_{\sigma^2}$  given by (3.4).

We obtain as in the proof of the first part, with a slight modification that the sequence  $(|\mathbb{T}_{n-1}|(\rho_n - \rho)/b_{|\mathbb{T}_{n-1}|})$  satisfies a MDP with speed  $b_{|\mathbb{T}_{n-1}|}^2$  and rate function  $I_\rho$  given by (3.5).

**5.4. Proof of Theorem 3.9.** Here also the proof works for the two cases.

Let us first deal with  $\hat{\sigma}_n$ . We have

$$\hat{\sigma}_n^2 - \sigma^2 = (\hat{\sigma}_n^2 - \sigma_n^2) + (\sigma_n^2 - \sigma^2).$$

From (4.10) and (4.32), we easily deduce that  $\sigma_n^2 \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} \sigma^2$  in the case 1 and in the case 2. Thus, it is enough to prove that  $\hat{\sigma}_n^2 - \sigma_n^2 \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0$ . Let  $\theta^{(0)} = (a_0, a_1, \dots, a_p)^t$ ,

$\theta^{(1)} = (b_0, b_1, \dots, b_p)^t$ ,  $\hat{\theta}_n^{(0)} = (\hat{a}_{0,n}, \hat{a}_{1,n}, \dots, \hat{a}_{p,n})$ ,  $\hat{\theta}_n^{(1)} = (\hat{b}_{0,n}, \hat{b}_{1,n}, \dots, \hat{b}_{p,n})$ .

Let us introduce the following function  $f$  defined for  $x$  and  $z$  in  $\mathbb{R}^{p+1}$  by

$$f(x, z) = \left( x_1 - z_1 - \sum_{i=2}^{p+1} z_i x_i \right)^2,$$

where  $x_i$  and  $z_i$  denote respectively the  $i$ -th component of  $x$  and  $z$ . One can observe that

$$\begin{aligned}\hat{\sigma}_n^2 - \sigma_n^2 &= \frac{1}{2|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} \left\{ f\left(\mathbb{X}_{2k}, \hat{\theta}_n^{(0)}\right) - f\left(\mathbb{X}_{2k}, \theta^{(0)}\right) \right\} \\ &\quad + \frac{1}{2|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} \left\{ f\left(\mathbb{X}_{2k+1}, \hat{\theta}_n^{(1)}\right) - f\left(\mathbb{X}_{2k+1}, \theta^{(1)}\right) \right\}.\end{aligned}$$

By the Taylor-Lagrange formula,  $\forall x \in \mathbb{R}^{p+1}$  and  $\forall z, z' \in \mathbb{R}^{p+1}$ , one can find  $\lambda \in (0, 1)$  such that

$$f(x, z') - f(x, z) = \sum_{j=1}^{p+1} (z'_j - z_j) \partial_{z_j} f(x, z + \lambda(z' - z)).$$

Let the function  $g$  defined by

$$g(x, z) = x_1 - z_1 - \sum_{j=2}^{p+1} z_j x_j.$$

Observing that

$$\begin{cases} \frac{\partial f}{\partial z_1}(x, z) = -2g(x, z) \\ \frac{\partial f}{\partial z_j}(x, z) = -2x_j g(x, z) \quad \forall j \geq 2, \end{cases}$$

we get easily that  $\left| \frac{\partial f}{\partial z_j}(x, z) \right| \leq 4(1 + \|z\|)(1 + \|x\|^2)$  for all  $j \geq 1$ , and this implies

$$|f(x, z') - f(x, z)| \leq c \|z' - z\| (1 + \|z\| + \|z' - z\|) (1 + \|x\|^2),$$

for some positive constant  $c$ . Now, applying the foregoing to  $f\left(\mathbb{X}_{2k}, \hat{\theta}_n^{(0)}\right) - f\left(\mathbb{X}_{2k}, \theta^{(0)}\right)$  and to  $f\left(\mathbb{X}_{2k+1}, \hat{\theta}_n^{(1)}\right) - f\left(\mathbb{X}_{2k+1}, \theta^{(1)}\right)$ , we deduce easily that

$$|\hat{\sigma}_n^2 - \sigma_n^2| \leq c \|\hat{\theta}_n - \theta\| \left(1 + \|\theta\| + \|\hat{\theta}_n - \theta\|\right) \frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} (1 + \|\mathbb{X}_k\|^2),$$

for some positive constant  $c$ . From the MDP of  $\hat{\theta}_n - \theta$ , we infer that

$$\|\hat{\theta}_n - \theta\| \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0. \quad (5.19)$$

Form Proposition 4.4 we deduce that

$$\frac{1}{|\mathbb{T}_{n-1}|} \sum_{k \in \mathbb{T}_{n-1, p-1}} (1 + \|\mathbb{X}_k\|^2) \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 1 + \text{Tr}(\Lambda). \quad (5.20)$$

We thus conclude via (5.19) and (5.20) that

$$\hat{\sigma}_n^2 - \sigma_n^2 \xrightarrow[b_{|\mathbb{T}_{n-1}|}^2]{\text{superexp}} 0.$$

This ends the proof for  $\hat{\sigma}_n$ . The proof for  $\hat{\rho}_n$  is very similar and uses hypothesis **(G2')** and **(N2')** to get inequalities similar to (4.10) and (4.32).

**Acknowledgments.** The authors thank Arnaud Guillin for all his advices and suggestions during the preparation of this work.

## REFERENCES

- [1] I. V. Basawa and J. Zhou. Non-Gaussian bifurcating models and quasi-likelihood estimation. *J. Appl. Probab.* 41A (2004), 55-64.
- [2] I. V. Basawa and R. M. Huggins. Extensions of the bifurcating autoregressive model for cell lineage studies. *J. Appl. Probab.* 36, 4 (1999), 1225-1233.
- [3] I. V. Basawa. and R. M. Huggins. Inference for the extended bifurcating autoregressive model for cell lineage studies. *Aust. N. Z. J. Stat.* 42, 4 (2000), 423-432.
- [4] I. V. Basawa, S. Y. Hwang and I. K. Yeo. Local asymptotic normality for bifurcating autoregressive processes and related asymptotic inference. *Statistical Methodology* 6 (2009), 61-69.
- [5] I. V. Basawa. J. Zhou. Least-squares estimation for bifurcating autoregressive processes. *Statist. Probab. Lett.* 74, 1 (2005), 77-88.
- [6] I. W. Basawa and J. Zhou. Maximum likelihood estimation for a first-order bifurcating autoregressive process with exponential errors. *J. Time Ser. Anal.* 26, 6 (2005), 825-842.
- [7] B. Bercu, B. de Saporta and A. Gégout-Petit. Asymtotic analysis for bifurcating autoregressive processes via martingale approach. *Electron. J. Probab.* 14 (2009), no. 87, 2492-2526 .
- [8] B. Bercu and A. Touati. Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab.* 18 (2008), no. 5, 1848-1869
- [9] V. Bitseki Penda, H. Djellout and A. Guillin. Deviation inequalities, Moderate deviations and some limit theorems for bifurcating Markov chains with application. *arXiv:1111.7303*
- [10] R. Cowan and R. G. Staudte. The bifurcating autoregressive model in cell lineage studies. *Biometrics* 42 (1986), 769-783.
- [11] J. F. Delmas and L. Marsalle. Detection of cellular aging in a Galton-Watson process. *Stoch. Process. and Appl.*, 120 : 2495-2519, 2010.
- [12] A. Dembo. Moderate deviations for martingales with bounded jumps. *Electron. Comm. Probab.* 1 (1996), no. 3, 11-17.
- [13] A. Dembo and O. Zeitouni. (1998) Large Deviations Techniques and Applications, 2nd Ed. (Springer, New York).
- [14] H. Djellout, A. Guillin and L. Wu. Moderate deviations of empirical periodogram and non-linear functionals of moving average processes. *Ann. Inst. H. Poincaré Probab. Statist.* 42 (2006), no. 4, 393-416.
- [15] H. Djellout. Moderate deviations for martingale differences and applications to  $\phi$ -mixing sequences *Stoch. Stoch. Rep.* 73 (2002),1-2, 37-63.
- [16] H. Djellout and A. Guillin. Large and moderate deviations for moving average processes. *Ann. Fac. Sci. Toulouse Math. (6)* 10 (2001), no. 1, 23-31.
- [17] N. Gozlan and C. Lonard. A large deviation approach to some transportation cost inequalities. *Probab. Theory Related Fields* 139 (2007), no. 1-2, 235-283.
- [18] N. Gozlan. Integral criteria for transportation-cost inequalities. *Electron. Comm. Probab.* 11 (2006), 64-77.
- [19] J. Guyon. Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab.* 17 (2007), no. 5-6, 1538-1569.
- [20] M. Ledoux. The concentration of measure phenomenon. *Mathematical Surveys and Monographs*, 89. American Mathematical Society, Providence, RI, 2001.
- [21] P. Massart. Concentration inequalities and model selection. *Lecture Notes in Mathematics*, 1896. Springer, Berlin, 2007.
- [22] A. Puhalskii. Large deviations of semimartingales: a maxingale problem approach. I. Limits as solutions to a maxingale problem. *Stochastics Stochastics Rep.* 61 (1997), no. 3-4, 141-243.
- [23] V. H. de la Peña; T. L. Lai and Qi-Man Shao. Self-normalized processes. Limit theory and statistical applications. *Probability and its Applications (New York)*. Springer-Verlag, Berlin, 2009. 275 pp.
- [24] B. de Saporta, A. Gégout-Petit and L. Marsalle Parameters estimation for asymmetric bifurcating autoregressive processes with missing data. *Electronic Journal of Statistics*, vol. 5 (2011) 1313-1353.
- [25] B. de Saporta, A. Gégout-Petit and L. Marsalle Asymmetry tests for Bifurcating Auto-Regressive Processes with missing data *arXiv:1112.3745*.



- [26] C. Z. Wei. Adaptive prediction by least squares predictors in stochastic regression models with applications to time series. *Ann. Statist.* 15, 4 (1987), 1667-1682.
- [27] J. Worms. Moderate deviations of some dependent variables. I. Martingales. *Math. Methods Statist.* 10 (2001), no. 1, 38-72.
- [28] J. Worms. Moderate deviations of some dependent variables. II. Some kernel estimators. *Math. Methods Statist.* 10 (2001), no. 2, 161-193.
- [29] J. Worms. *Principes de déviations modérées pour des martingales et applications statistiques*. Thèse de Doctorat à l'Université Marne-la-Vallée, 2000.
- [30] J. Worms. Moderate deviations for stable Markov chains and regression models. *Electron. J. Probab.* 4 (1999), no. 8, 28 pp.

*E-mail address:* Valere.Bitsekipenda@math.univ-bpclermont.fr

LABORATOIRE DE MATHÉMATIQUES, CNRS UMR 6620, UNIVERSITÉ BLAISE PASCAL, AVENUE DES LANDAIS, 63177 AUBIÈRE, FRANCE.

*E-mail address:* Hacene.Djellout@math.univ-bpclermont.fr

LABORATOIRE DE MATHÉMATIQUES, CNRS UMR 6620, UNIVERSITÉ BLAISE PASCAL, AVENUE DES LANDAIS, 63177 AUBIÈRE, FRANCE.